

**VERBESSERTER SCHÄTZANSÄTZE BEI KONTROLLIERTER FEHLSPEZIFIKATION
ÖKONOMETRISCHER GLEICHUNGSSYSTEME, DEREN MODELLKONFORME
INTERPRETATION SOWIE EIN TEST ZUR ERWEITERTEN ANWENDUNG
VON ORDINARY LEAST SQUARES AUF INTERDEPENDENTE
GLEICHUNGSSYSTEME**

**Inaugural-Dissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschafts- und Sozialwissenschaften
(Dr. rer. pol.)**

**an dem Fachbereich Wirtschafts- und Sozialwissenschaften
der Friedrich-Alexander-Universität Erlangen-Nürnberg**

**vorgelegt von
Diplom-Kaufmann Dieter Drefahl
aus Badenweiler**

1980

VERBESSERTE SCHÄTZANSÄTZE BEI KONTROLLIERTER FEHLSPEZIFIKATION
ÖKONOMETRISCHER GLEICHUNGSSYSTEME, DEREN MODELLKONFORME INTER-
PRETATION SOWIE EIN TEST ZUR ERWEITERTEN ANWENDUNG VON ORDINARY
LEAST SQUARES AUF INTERDEPENDENTE GLEICHUNGSSYSTEME

Inaugural-Dissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschafts- und Sozialwissenschaften
(Dr. rer. pol.)

an dem Fachbereich Wirtschafts- und Sozialwissenschaften
der Friedrich-Alexander-Universität Erlangen-Nürnberg



vorgelegt von

Diplom-Kaufmann Dieter D r e f a h l
aus Badenweiler



Referent : Prof. Dr. H. Schneeberger

Korreferent : Prof. Dr. M. Neumann

Promotion : 15. Juli 1980

Meinen Eltern

gewidmet

Vorwort

Die vorliegende Arbeit entstand am Lehrstuhl für Statistik II der Wirtschafts- und Sozialwissenschaftlichen Fakultät in Nürnberg. Während einer Zeit von acht Jahren wurde die wissenschaftliche Denk- und Arbeitsweise des Verfassers vom Inhaber dieses Lehrstuhls, Prof. Dr. H. Schneeberger, zuerst bei dem Studenten geweckt, dann bei dem Assistenten geprägt und schließlich bei dem Doktoranden auf ein Ziel gerichtet. So behutsam diese Anleitung und so groß der bestimmt außergewöhnliche wissenschaftliche Freiraum auch waren - die Entwicklung auf dem schmalen Grat zwischen formalistischen Selbstzweck einerseits und formal nicht abgesicherter Spekulation andererseits hätte ohne diese "unsichtbare Hand" wohl kaum stattfinden können. So muß der Verfasser für mögliche Fehler dieser Arbeit selbstverständlich verantwortlich zeichnen, während er hoffentlich vorhandene positive Elemente oder Ergebnisse nicht für sich allein beanspruchen kann, ohne voller Dankbarkeit die Schule zu nennen.

Neben den fruchtbaren Diskussionen mit den Kollegen Dipl. Kfm. J. Pöllot und Dipl. Math. Dr. rer. pol. W. Goller halfen vor allem die aus praktischer Erfahrung in der Anwendung ökonometrischer Methoden gewachsenen Anregungen von Dipl. Volksw. Dipl. Kfm. Dr. rer. pol. A. Haid. Dieser war auch das unermüdliche Bindeglied zu dem Lehrstuhl für Volkswirtschaftslehre, dessen Inhaber Prof. Dr. M. Neumann der Verfasser neben anderem besonders dafür zu danken hat, daß er sich nicht nur zur Korrektur eines derart spröden Themas bereit erklärte, sondern auch die für den Verfasser wichtigen Bezüge zur Volkswirtschaftstheorie aufzeigte.

Nürnberg, im Juli 1980



Digitized by the Internet Archive
in 2026

I N H A L T S V E R Z E I C H N I S

= = = = =

Vorbemerkungen

1.	<u>Die Annahmen der linearen Regression</u>	S. 1
1.1.	Eingleichungsregression mit nichtstochastischen ("fixen") Regressoren	S. 1
1.2.	Eingleichungsregression mit stochastischen Regressoren	S. 9
2.	<u>Verletzung der Modellannahmen des OLS - Schätzers durch falsche Wahl der Regressorenmenge</u>	S. 15
2.1.	Definition der Fehlertypen	S. 16
2.1.1.	Fehlspezifikation nach Fehlertyp I	S. 17
2.1.2.	Fehlspezifikation nach Fehlertyp II	S. 18
2.1.3.	Fehlspezifikation nach Fehlertyp III	S. 19
3.	<u>Verletzung der Modellannahmen des OLS - Schätzers durch Verwendung eines unzulässigen Regressortyps</u>	S. 20
3.1.	Einregressormodell mit fehlerbehaftetem Regressor	S. 21
3.2.	Mehrregressormodell mit einem fehlerbehafteten Regressor	S. 21
3.3.	Mehrregressormodell mit mehreren fehlerbehafteten Regressoren	S. 24
3.4.	Die Bedeutung fehlerbehafteter Regressoren	S. 25
3.4.1.	Proxy - Variable	S. 25
3.4.2.	Meßfehlerbehaftete Regressoren	S. 28
3.5.	Bekannte Schätzansätze für das "Fehler-in-den-Regressoren" - Modell (EIV-Modell)	S. 28
3.5.1.	Der Instrumentalvariablenansatz	S. 28
3.5.1.1.	Lagvariable als Instrumentalvariable	S. 31
3.5.1.2.	Klassenzugehörigkeit als Instrumentalvariable	S. 34

3.6.	Bewertung der Lösungsansätze	S. 39
3.7.	Kritische Modelle	S. 40
3.7.1.	Modelle mit verzögerten endogenen Variablen als Regressoren und autokorrelierten Fehlern	S. 40
3.7.2.	Durch Transformation in absolute Differenzen entstandene Modelle	S. 41
3.7.3.	Regelungsmodelle (Rückkopplungsmodelle)	S. 42
4.	<u>Gemeinsamkeiten von "unterspezifiziertem" und "Fehler-in-dem-Regressor"-Regressionsansatz</u>	S. 46
5.	<u>Seemingly unrelated regression - Systeme</u>	S. 51
5.1.	Die Koeffizientenschätzung im SUR-System nach Zellner	S. 52
6.	<u>Alternativer SURE-Ansatz</u>	S. 59
6.1.	SURE-Ansatz für das 'EIV' - Modell	S. 60
6.2.	Inkonsistenter SURE-Ansatz bei Unterspezifikation	S. 67
6.3.	Konsistenter SURE-Ansatz bei Unterspezifikation	S. 72
6.4.	Beispiele	
6.4.1.	Beispiel des Anhangs Seite A.V.	S. 77
6.4.2.	Beispiel des Anhangs Seite A.VI.	S. 77
6.4.3.	Beispiel des Anhangs Seite A.VII.	S. 79
6.4.4.	Beispiel des Anhangs Seite A.VIII.	S. 80
6.4.5.	Beispiel des Anhangs Seite A.IX.	S. 81
6.4.6.	Beispiel des Anhangs Seite A.X.	S. 84
7.	<u>Die Bedeutung der Struktur- und reduzierten Form des interdependenten ökonometrischen Gleichungssystems</u>	S. 97
7.1.	Ein erläuterndes Beispiel	S. 100
7.1.1.	Exakte Identifikation	S. 107

7.1.2.	Überidentifikation	S. 108
8.	<u>Die Aussagen interdependenter Gleichungs-</u> <u>systeme</u>	S. 120
8.1.	Das geschlossene Gleichgewichtssystem	S. 120
8.1.1.	Die Strukturform	S. 121
8.1.2.	Die reduzierte Form	S. 122
8.2.	Das offene "Kausalketten" - System	S. 124
8.2.1.	Rekursive Systeme	S. 124
8.2.2.	Nichtrekursive Systeme	S. 134
9.	<u>Der Test auf Kausalketten - Beziehungen</u>	S. 139
10.	<u>Schlußbemerkungen</u>	S. 146

Anhang (Numerische Beispiele)

Literaturverzeichnis

Vorbemerkungen

Statistische Verfahren werden in zunehmendem Maße im Rahmen ökonometrischer Untersuchungen zur empirischen Überprüfung wirtschaftstheoretischer Hypothesen eingesetzt. Diese Entwicklung wurde nicht zuletzt durch den Ausbau der statistischen Methodenbanken für elektronische Rechenanlagen gefördert, die selbst hochentwickelte statistische Methoden scheinbar problemlos für den Anwender zugänglich machen. Ein wesentliches Ziel der vorliegenden Arbeit ist es, die Gefahren einer kritiklosen Anwendung statistischer Verfahren auf reale Problemstellungen aufzuzeigen. Schon bei dem als "unkritisch" geltenden Schätzverfahren der multiplen Regression nach der Methode der kleinsten Quadrate ("OLS") wird in den ersten drei Kapiteln deutlich gemacht, daß die für dieses Schätzverfahren notwendigen Annahmen nur unter starkem Vorbehalt als realistisch akzeptiert werden können. Dabei werden vor allem zwei Ursachen für eine Verletzung dieser Annahmen bei einer wirtschafts- und sozialwissenschaftlichen Anwendung der Regression herausgearbeitet: Im Gegensatz zu den Naturwissenschaften mit ihren Möglichkeiten des kontrollierten Experiments ist es nach Meinung des Verfassers im wirtschafts- und sozialwissenschaftlichen Bereich geradezu typisch, daß unvollständige Schätzspezifikationen und fehlerbehaftete Regressoren ("Proxies") vorliegen. Nach einer Behandlung der Folgen einer solchen Verletzung der Schätzannahmen und einem kurzen Abriß über die bekannten Lösungsansätze dieser Schätzprobleme wird in Kapitel 4 ein abgestuftes "Informationsdefizit" des Ökonometrikers als gemeinsames Merkmal der angesprochenen Schätzprobleme formuliert. Nach der Vorstellung eines bekannten Schätzverfahrens ("SURE-Schätzer nach Zellner") in Kapitel 5, das nach Lehrbuch-

meinung eine Lösung für die angesprochenen Probleme liefert, werden in Kapitel 6 eigene Lösungsvorschläge vorgetragen, welche die Modellvorstellung des Zellner - Schätzers übernehmen, jedoch durch die Einführung zusätzlicher Annahmen im Gegensatz zum inkonsistenten Zellner - Schätzer zu konsistenten Schätzungen führen. Dies wird neben der formalen Darstellung der Schätzverfahren auch an konstruierten Simulationsrechnungen und an einem konkreten Beispiel aus der Literatur veranschaulicht. Da für diese Schätzansätze die zusätzlichen Informationen aus einem die zu schätzende Gleichung 'umgebenden', aber nicht interdependenten Gleichungssystem stammen, werden in Kapitel 7 die Modellvorstellungen interdependenter Gleichungssysteme untersucht, da die zusätzliche Interdependenz der einzelnen Gleichungen eine zusätzliche Information erhoffen ließ. Hierbei wurde jedoch festgestellt, daß eine Interdependenz lediglich durch eine komparativ-statische Modellvorstellung postuliert wird, aber in den bei interdependenten Systemen eingeführten Schätzverfahren keine "kausale" Verknüpfung der Einzelgleichungen nach sich zieht. Im gleichen Kapitel wird daraufhin anhand eines kleinen Beispiels gezeigt, wie eine Mißachtung dieser Grundlagen zu inkonsistenten Schätzungen und falschen Interpretationen führen kann. Nach der Behandlung der modelladäquaten Interpretation interdependenter Systeme wird in Kapitel 8 gezeigt, welche inhaltlichen Besonderheiten rekursive Systeme als Sonderfall interdependenter Systeme aufweisen. Durch eine neue Formulierung der für eine konsistente OLS - Schätzung notwendigen Bedingungen kann im gleichen Kapitel gezeigt werden, daß unter diesen Bedingungen auch in nichtrekursiven Modellen OLS konsistente Schätzungen liefert und wegen einer dann vorliegenden gerichteten "kausalen" Verknüpfung innerhalb und zwischen den

einzelnen Gleichungen zusätzliche Informationen für den Ökonometriker liefert. In Kapitel 9 schließlich wird ein Test dieser Bedingungen eingeführt.

Die verwendete Nomenklatur richtet sich weitgehend nach der von Johnston (1972) benutzten Schreibweise; Abweichungen in der Darstellung werden in den jeweiligen Kapiteln durch eine unterschiedliche inhaltliche Aussage begründet. Mit Ausnahme des Schätzverfahrens nach Zellner werden alle in dieser Arbeit angesprochenen Schätzverfahren für Ein- und Mehrgleichungssysteme nicht explizit dargestellt, da sie in allen ökonometrischen Lehrbüchern vorgestellt werden und als bekannt vorausgesetzt werden können.

§1. Die Annahmen der linearen Regression

=====

In der Regressionstheorie spielen die Residuen einer Regression eine ungemein wichtige Rolle: Erst durch eine genaue Definition ihrer Eigenschaften werden die in der Regressionstheorie verwendeten Tests möglich, und die Aussagen über Konsistenz bzw. Unverzerrtheit irgendeines Regressionsverfahrens hängen immer von der Erfüllung bestimmter Eigenschaften der Residuen ab. Diese Eigenschaften sollen anhand der Forderungen des wichtigsten Schätzverfahrens der Regression untersucht werden, nämlich der Methode der kleinsten Quadrate ¹⁾. Dabei wird besonderer Wert auf die Untersuchung der inhaltlichen Aussagen der formalen Forderungen gelegt, denn in der Ökonometrie ist der Untersuchungsgegenstand immer ein realer Prozess, der durch ein vereinfachtes Modell abgebildet werden soll.

Das einfachste Modell wird durch eine Eingleichungsregression dargestellt; Mehrgleichungssysteme bedeuten zusätzliche Implikationen, die in späteren Kapiteln untersucht werden. Die grundsätzlichen Überlegungen werden jedoch der Übersichtlichkeit wegen an Eingleichungsmodellen durchgeführt.

1.1. Eingleichungsregression mit nichtstochastischen ("fixen") Regressoren

In dem Modell

$$y_t = \eta_t + v_t = \sum_{i=1}^K x_{it} \gamma_i + v_t \Rightarrow \sum x_i \gamma_i + \underline{v} = X \underline{\gamma} + \underline{v}$$

seien $\underline{x}_i$ ($i=1, \dots, K$) die Spaltenvektoren der Länge T, die die Abweichungen vom jeweiligen arithmetischen Mittel der insgesamt T Messungen der unabhängigen Variablen (im folgenden "Regressoren"

1) In der weiteren Arbeit wird hierfür die Abkürzung OLS (Ordinary Least Squares) verwendet.

genannt) darstellen mit $t=1, \dots, T$,

$\underline{y}$ sei der Spaltenvektor der Länge T, der die Abweichungen der abhängigen Variablen (im folgenden "Regressand" genannt) vom arithmetischen Mittel aller ihrer T Messungen darstellt,

$\underline{\gamma}$ sei der Spaltenvektor der Länge K, der die unbekannten Regressionskoeffizienten der insgesamt K Regressoren darstellt. Dabei seien die γ_i ($i=1, \dots, K$) konstante, d. h. nichtstochastische Koeffizienten ¹⁾.

$\underline{v}$ sei der Spaltenvektor der Länge T, der die nicht meßbaren, zufälligen Bestandteile des Modells enthält.

Diese Bestandteile, die im folgenden kurz "Residuen" genannt werden, machen eine über das bloße Auflösen eines linearen Gleichungssystems hinausgehende "Ausgleichsrechnung" nötig, wenn die Koeffizienten γ_i geschätzt ²⁾ werden sollen. Für diese Schätzung werden bezüglich der v_t bestimmte Annahmen getroffen, die neben der statistisch-formalen Notwendigkeit auch interpretatorische Aussagen beinhalten. Diese Annahmen sind:

Annahme 1: $E(v_t) = 0$

Das heißt, der Erwartungswert des Residuums ist für einen bestimmten Zeitpunkt t Null. Formal ist diese Annahme nötig, um eine Unterscheidung zwischen der Konstanten ³⁾ und dem Residuum zu ermöglichen. Für die Interpretation der zu schätzenden Gleichung bedeutet das: Neben den in η_t zusammengefaßten Einflußgrößen existieren weitere, den Meßwert von y_t bestimmende

1) Faßt man die γ_i als Zufallsgrößen auf, wird eine Analyse wesentlich erschwert, aber nicht unmöglich; vgl. z. B. Maddala(1977) S. 390 ff.

2) Unter "Schätzung" werden in diesem Kapitel immer OLS-Schätzungen verstanden, soweit nicht ausdrücklich ein anderes Schätzverfahren angesprochen wird.

3) Die in dem dargestellten Modell durch die Verwendung der Mittelwertabweichungen der Variablen nicht auftaucht!

Einflußfaktoren, deren Gesamteinfluß im Erwartungswert jedoch verschwindet. Dabei kann man den Gesamtfehler v_t in zwei unterschiedliche Bestandteile aufspalten ¹⁾:

- a) Meßfehler, die beim Messen ²⁾ von y_t ³⁾ begangen werden.

Hier verlangt die Annahme $E(v_t)=0$ das Fehlen von systematischen Verzerrungen bei der Messung. Dieser Annahme wird zwar von Praktikern widersprochen ⁴⁾, sie kann aber bei entsprechend vorsichtiger Interpretation des geschätzten konstanten Terms postuliert werden.

- b) Zufallsfehler, die auf der grundsätzlichen Unmöglichkeit der Wiederholung einer Messung unter exakt identischen Bedingungen beruhen. Diese Zufallsfehler sind die Summe aller Einflußgrößen, die nicht durch die x_{it} erfaßt werden. Dies sind:

- b₁) Als Einflußgrößen bekannte, aber nicht meßbare Größen wie z. B. Stimmungen, Erwartungshaltungen. Dabei liegt fehlende Meßbarkeit auch dann vor, wenn das Skalenniveau der Messung für eine Regression nicht geeignet ist wie z. B. bei ordinalskalierten Daten.
- b₂) Einmalige oder sehr selten bei einer Messung auftretende Einflüsse wie z. B. Naturkatastrophen, die wegen der Seltenheit ihres Auftretens nicht als Einflußgrößen erkannt und in den Modellansatz einbezogen werden (z. B. als Dummyvariablen).
- b₃) Nicht bekannte Einflußgrößen, deren Einfluß über den einer 'seltenen Störung' in ihrer Wirkungsdauer hinausgeht.

1) Vgl. z. B. Wonnacott & Wonnacott (1970) S. 17

2) "Messen" muß nicht unmittelbares Bestimmen der Meßgröße bedeuten, sondern kann auch (gerade in ökonometrischen Modellen) ein indirektes Herleiten aus anderen Größen (z. B. in der gesamtwirtschaftlichen Kontenrechnung) oder eine hochgerechnete Stichprobe (z. B. bei den Investitionen) bedeuten.

3) Meßfehler der Regressoren werden in Kapitel 3 behandelt

4) Vgl. z. B. Stobbe (1972) S. 54 oder
Esenwein-Rothe (1972) S. 499 ff.

Die Forderung, daß der Erwartungswert der Summe dieser Bestandteile Null ist, wird bei konkreten Modellen wohl nicht immer erfüllt. Da eine Verletzung dieser Forderung jedoch lediglich eine Fehlschätzung des konstanten Terms nach sich zieht und dieser Term oft überhaupt nicht oder nur mit vorsichtigen Vorbehalten interpretiert wird, ist die Annahme 1 sicher für die meisten Anwendungen der Regression postulierbar. Sie kann nicht getestet werden, sondern muß aus a-priori-Gründen angenommen werden ¹⁾.

Das Auftreten eines Residuums in einer Regression bedeutet also, daß entweder die Größe y_t fehlerhaft gemessen oder Einflußgrößen ausgeschlossen wurden. Ob dadurch neben der möglichen Verzerrung des konstanten Terms die meist mehr interessierenden Regressionskoeffizienten γ_i ($i=1, \dots, K$) dadurch verzerrt geschätzt werden, hängt von der Erfüllung der folgenden Annahme ab.

Annahme 2: $E(v_t/x_{it}) = 0$ für alle i und t ²⁾

Formal ist diese Annahme nötig, um eine Inkonsistenz der Schätzungen der γ_i zu vermeiden ³⁾. Sie ist damit eine der wichtigsten Annahmen der Regressionsrechnung und muß deshalb bei jeder praktischen Anwendung sorgfältig auf ihre Erfüllung geprüft werden. Deshalb ist es überraschend, daß diese Prüfung bei der Veröffentlichung von Regressionsschätzungen so gut wie nie Bestandteil der Veröffentlichung ist. Bedenklich wird dies besonders dann, wenn eine geschätzte Regression ein kleines Bestimmtheitsmaß ⁴⁾ liefert. Denn nach den auf Seite 3 gemachten Ausführungen kann dies auf einem großen Fehleranteil ⁵⁾ bei der Messung von y oder auf einem großen Anteil der nicht im Regressionsansatz

1) Vgl. Koutsoyiannis (1973) S. 175

2) Annahme 2 schließt Annahme 1 ein; eine Analyse der inhaltlichen Aussagen wird aber bei isolierter Betrachtung deutlicher.

3) Vgl. z. B. Maddala (1977) S. 152 f.

4) Zur Definition des Bestimmtheitsmaßes vgl. z. B. Schneeweiß (1974) S. 128

5) Unter "Anteilen" sind hier immer Varianzanteile zu verstehen

enthaltenen Variablen beruhen. Für den Meßfehler ist die Annahme 2 wohl im allgemeinen mit der Art seiner Entstehung ¹⁾ zu begründen, z. B. bei einer Hochrechnung aus einer zwar mit einer Fehlervarianz behafteten, aber unverzerrten Stichprobe. Vorsicht ist jedoch geboten, wenn das Bestimmtheitsmaß klein ist und gleichzeitig von dem Meßfehler bekannt ist, daß er gering ist. Denn dann weiß man, daß man Variablen nicht in den Schätzansatz aufgenommen hat, die einen nicht unwesentlichen Anteil von y erklären ²⁾. Und mit Ausnahme des unter b_1 aufgeführten Falles gibt es keinen ersichtlichen Grund, warum man ohne sorgfältige Prüfung von Einflußgrößen ausgehen kann, von denen man zwar nichts weiß, deren Unkorreliertheit mit den im Schätzansatz auftretenden Regressoren man aber zur Erzielung unverzerrter Schätzungen postulieren muß!

Im allgemeinen jedoch verteidigt man diese Annahme mit folgendem Argument ³⁾: Die Größe v_t ist die Überlagerung sehr vieler, aber in ihrem Einzeleinfluß unbedeutender Restvariablen, die zudem noch von Messung zu Messung wechseln. Und erst diese Überlagerung ist in ihrer Summation im Erwartungswert unkorreliert mit den Regressoren des Schätzansatzes ⁴⁾. Damit soll vor allem angedeutet werden, daß die "Vernachlässigung" einer oder einiger weniger Variablen mit hohem partiellem Einfluß ausgeschlossen wird, wohl in der - bisweilen sehr optimistischen - Annahme, daß solche Variablen mit einem hohen Varianzbeitrag von dem Ökonometriker erkannt würden. Betrachtet man jedoch den gerade in der Ökonometrie häufig auftretenden Fall einer

1) Vgl. Fußnoten 2) und 4) der S. 3

2) So bezeichnet Theil ((1971) S. 104) generell das Residuum v als "Kombination der Effekte von vernachlässigten Variablen".

3) Vgl. z. B. Schneeweiß (1974) S. 39

4) Mit derselben Argumentation wird ja auch die Annahme der Normalverteilung der Residuen begründet, die für praktisch alle Tests der Regression und für die Herleitung der Maximum-Likelihood-Ansätze benötigt wird.

Regression von Zeitreihenwerten, so müßten die vernachlässigten Variablen mit dem Trend unkorreliert sein; diese Annahme ist aber zumindest dann schwer zu begründen, wenn die vernachlässigten Restvariablen aus dem gleichen Bereich stammen wie die Regressoren des Modellansatzes ¹⁾. Wenn allerdings ein hohes Bestimmtheitsmaß gefunden wird, so kann man selbst bei hoher Korrelation zwischen $\underline{y}$ und den $\underline{x}_i$ wegen des "proximity theorems" von Wold ²⁾ die Verzerrung der geschätzten Koeffizienten $\hat{\gamma}_i$ vernachlässigen ³⁾. Es darf hierbei jedoch nicht übersehen werden, daß ein hohes Bestimmtheitsmaß nicht unbedingt die Bestätigung einer richtigen Spezifikation ist. Gerade bei Zeitreihendaten findet man im allgemeinen eine hohe Korrelation zwischen willkürlich gewählten Variablen, wenn sie nur einem ähnlichen Datenbereich entstammen. Und da es der Ökonometriker ja mit Wirtschaftsdaten zu tun hat und es kaum Bereiche gibt, die losgelöst vom - zumindest makroökonomischen - Wirtschaftsgeschehen sind, ist die Gefahr von "Nonsens-Korrelationen" sehr groß ⁴⁾.

In allen folgenden Abschnitten dieser Arbeit wird jedoch davon ausgegangen, daß zumindest die Spezifikation der K Regressoren durch (z. B. volkswirtschaftliche) Hypothesen begründet und auch richtig ist, d. h. daß sowohl ihre Zusammenstellung wie auch ihre Funktionalform (linear, nichtlinear in den Variablen) zumindest eine Teilmenge der $\underline{x}_i$ des "wahren" Modells ist. Auch die Annahme 2 kann formal nicht getestet werden ⁵⁾.

Annahme 3: $E(\underline{v}\underline{v}'/X) = \sigma^2(v) \cdot I$, wobei I die T·T-Einheitsmatrix ist. Diese Annahme ist formal nötig für den Nachweis der

1) Z. B. ebenfalls konjunkturabhängige makroökonomische Variablen

2) Wold u. Faxer (1957)

3) Ein Beispiel hierfür liefert Maddala (1977) S. 153 für $K = 1$

4) Beispiele findet man z. B. bei Wonnacott & Wonnacott (1970) S. 125

5) Vgl. Koutsoyiannis (1973) S. 173

BLUE-('best linear unbiased estimator' ¹⁾) Eigenschaft des OLS-Schätzers. Bei Verletzung dieser Annahme ist der OLS-Schätzer wegen Autokorrelation oder Heteroskedastizität der Residuen ²⁾ zwar noch immer unverzerrt, aber nicht mehr BLUE ³⁾. Außerdem sind die Schätzungen für $\sigma^2(v)$ verzerrt und damit auch die auf $\sigma^2(v)$ aufbauenden Tests ³⁾. Solange jedoch nicht mindestens die Konsistenz eines Schätzers gesichert ist, scheint es dem Verfasser müßig, sich über seine Effizienz auszulassen. Deshalb wird auch der Schwerpunkt der folgenden Kapiteln auf Konsistenzuntersuchungen liegen und die Frage der Effizienz erst nachgeordnet diskutiert. Dabei wird unter den unterschiedlichen Definitionen der Konsistenz die von Wonnacott & Wonnacott ⁴⁾ gewählt: " $\hat{\theta}$ ist dann und nur dann ein konsistenter Schätzer, wenn sowohl seine Verzerrung wie seine Varianz gegen Null gehen, wenn der Stichprobenumfang gleich dem Umfang der Gesamtheit ist". Dies bedeutet für den Fall der Regression, daß T gegen Unendlich geht. Diese Definition ist z. B. für Maddala ⁵⁾ und Johnston ⁶⁾ lediglich eine hinreichende, jedoch keine notwendige Bedingung für die Konsistenz. Diese Definition hat jedoch aus der Sicht des Verfassers zwei Nachteile. Erstens wären mit dieser Definition die Konstruktion von Verteilungen möglich, für die zwar erwartungstreue Schätzungen gebildet werden können, die Bildung von 'konsistenten' Schätzern aber

1) Vgl. z. B. Schneeweiß (1974) S. 72 f. Die Eigenschaft "best" bedeutet allerdings nicht, daß OLS der "beste" Schätzer überhaupt ist, sondern nur die kleinstmögliche Varianz besitzt. Hier ist also ebenso wie bei der Herleitung des OLS-Schätzers selbst die Verlustfunktion ('loss function') von quadratischer Form. Daß die Wahl dieser Verlustfunktion nicht unumstritten ist, kann im Rahmen dieser Arbeit nur erwähnt, aber nicht diskutiert werden.

2) Über die Eigenschaften des OLS-Schätzers bei Auftreten von Autokorrelation und Heteroskedastizität der Residuen ist noch wenig bekannt. Vgl. Schneeweiß (1974) S. 194

3) Vgl. z. B. Maddala (1977) S. 259 und S. 283

4) Wonnacott & Wonnacott (1970) S. 46

5) Maddala (1977) S. 148

6) Johnston (1972) S. 272

als unmöglich bewiesen werden kann ¹⁾. Umgekehrt aber wäre die Existenz von 'konsistenten', aber nicht asymptotisch unverzerrten Schätzern möglich ²⁾. Zweitens wird gerade in der Regressionstheorie die Konsistenz fast immer über die Eigenschaft asymptotischer Unverzerrtheit und verschwindender Varianz geführt, so daß eine Unterscheidung zwischen diesen Eigenschaften und der 'Konsistenz' keinen Vorteil verspricht. Damit erscheint dem Verfasser die Definition nach Wonnacott & Wonnacott die in ihrer Auswirkung anschaulichere zu sein und wird deshalb bei den weiteren Ausführungen verwendet; wenn zur Frage der Konsistenz eines Schätzers ein Autor mit abweichender Definition zitiert wird, so geschieht dies selbstverständlich nur dann, wenn er zum Beweis den Nachweis der beiden Eigenschaften benutzt oder ausdrücklich ihre Existenz zeigt.

Annahme 4: Die Matrix $X'X$ sei nichtsingulär.

Diese Annahme ist sofort nachprüfbar, da ja diese Matrix keine stochastischen Elemente enthält und ihr Rang sofort nachprüfbar ist. Formal ist diese Annahme nicht zur Bestimmung von $\hat{\eta}$, wohl aber zur eindeutigen Schätzung der einzelnen Regressionskoeffizienten $\hat{\gamma}_1$ nötig. Eine inhaltliche Interpretation erübrigt sich damit, da ja keine Plausibilitätsüberlegungen an die Stelle nicht nachprüfbarer Annahmen treten müssen.

Annahme 5: Die Residuen v_t sind bei gegebenem X unabhängig von t identisch normalverteilt

Formal ist diese Annahme zur Anwendung des größten Teils der bei der Regressionsanalyse benutzten Tests nötig. Begründet wird diese Annahme über den zentralen Grenzwertsatz, wobei

1) Gabrielsen (1978)

2) Johnston (1972) S. 272 und Intriligator (1978) S. 116

für die Annahme der "sehr großen Zahl" der die Normalverteilung bildenden Summanden die schon auf Seite 5 angeführten Argumente benutzt werden. Somit kann sich eine Plausibilitätsprüfung auf die dort aufgeführten Überlegungen stützen.

Die Annahme 5 kann statistisch getestet werden; allerdings müssen für diese Tests die Annahmen 1 bis 4 erfüllt sein.

Die gebräuchlichsten Tests sind z. B. bei Maddala ¹⁾ aufgeführt.

1.2. Eingleichungsregression mit stochastischen Regressoren

In dem jetzt vorliegenden Modell

$$\underline{Y} = \sum_{i=1}^K \underline{x}_i \gamma_i + \underline{v} = \underline{X} \underline{\gamma} + \underline{v}$$

seien in der Matrix $\underline{X}$ insgesamt K_s stochastische Regressoren $\underline{x}_{(s)}$ enthalten mit $1 \leq K_s \leq K$. Das heißt, mindestens ein Paar $\underline{x}_{(s)}; \underline{v}$ hat eine bivariate Verteilung ²⁾. Dieser Fall liegt bei ökonometrischen Untersuchungen sicher häufiger vor als der Fall ausschließlich fixer Regressoren, da "die Beobachtungen selten durch ein kontrolliertes Experiment erzeugt werden" ³⁾. Zu beachten ist jedoch, daß die zufälligen Ausprägungen der $\underline{x}_{(s)}$ ohne Fehler gemessen werden. Typische Beispiele für stochastische Regressoren sind klimatische Daten (Regenmenge, Sonnendauer), das Steueraufkommen (nicht der Steuersatz!), die umlaufende Geldmenge und, als typische Gruppe, alle verzögerten endogenen Variablen.

Es kann gezeigt werden ⁴⁾, daß bei 'unabhängiger stochastischer linearer Regression', d. h. vor allem mit $E(\underline{v}/\underline{X}) = 0$,

1) Maddala (1977) S. 306 f.

2) Vgl. z. B. Malinvaud (1970) S. 100 f.

3) Maddala (1977) S. 148

4) Vgl. Schönfeld (1971) Band II S. 19 ff., vor allem auch S. 21 Satz 8.4/1. oder Maddala (1977) S. 151 f.

OLS-Schätzungen unverzerrt sind ¹⁾. Selbst bei der abgeschwächten Annahme der 'kontemporär unkorrelierten stochastischen linearen Regression', nämlich $\text{plim}(\frac{1}{T} X'v) = \emptyset$, ist die Konsistenz der OLS-Schätzungen gesichert ²⁾. Deshalb soll im Folgenden nur diese schwächere Annahme untersucht werden.

Annahme 1_(s): $E(v_t) = 0$

Hier gilt das bereits in Abschnitt 1.1. Gesagte.

Annahme 2_(s): $\text{plim}(\frac{1}{T} X'v) = \emptyset$

Formal ist diese Bedingung wieder zur Erreichung der Konsistenz nötig. Inhaltlich treten wieder die Probleme des Abschnitts 1.1 (Annahme 2) auf, wobei hier jedoch noch die zusätzliche Implikation der Zufälligkeit der Regressoren auftritt. Denn jetzt muß zusätzlich noch darauf geachtet werden, daß die Verteilung dieser Zufallsgrößen mit der Verteilung der Residuen zumindest kontemporär unkorreliert ist. Diese Annahme ist bei dem Beispiel der klimatischen Daten - zumindest nach den heutigen technischen Möglichkeiten - sicher erfüllt ³⁾. Bei Wirtschaftsdaten, die aus dem gleichen Datenbereich stammen, ist diese Annahme jedoch sorgfältig zu prüfen. Betrachtet man z. B. den Fall "Der Staat plant eine bestimmte Höhe von Staatsausgaben ⁴⁾, die Ausgaben am Ende des Jahres stimmen aber nicht mit den Plandaten überein", so hat man den typischen Fall einer (fehlerfrei gemessenen!) Zufallsgröße, die geradezu als Prototyp einer exogenen Steuergröße betrachtet werden kann ⁵⁾. Annahme 2_(s)

1) Wenn auch nicht mehr BLUE, da die Schätzfunktion für $\hat{\gamma}$ nicht mehr linear in den Zufallsvariablen ist; vgl. Schönfeld (1971) Band II S. 4 Satz 8.1.3

2) Vgl. Schönfeld (1971) Band II S. 24 Satz 8.5/1.

3) "Die Zufallsverteilung der Regenfälle ist mit als Regressoren auftretenden Wirtschaftsgrößen nicht korreliert"

4) Auf das Problem, daß Steuergrößen in einem Regelkreis von v abhängig werden können, wird in Kapitel 4 eingegangen.

5) So z. B. von Schneeweiß (1974) S. 244

kann aber nur dann gelten, wenn die Verteilung derjenigen Einflüsse, die die Abweichungen der tatsächlichen Ausgaben von den geplanten Ausgaben verursachten, mit der Verteilung der Residuen v unkorreliert ist - eine Annahme, die nur dann plausibel ist, wenn diese Einflüsse nicht aus dem wirtschaftlichen Bereich stammen.

Für das Modell mit stochastischen Regressoren ist man aber nicht nur auf Plausibilitätsüberlegungen angewiesen. Hier hat man die scheinbare Hilfe von Tests, die die Unkorreliertheit der Verteilungen statistisch überprüfen lassen ¹⁾. Scheinbar ist diese Hilfe jedoch deshalb, weil diese Tests die Existenz von mindestens $K_S + 1$ nichtstochastischen Instrumentalvariablen ²⁾ fordern. Diese werden in der ökonometrischen Praxis aber nur in Ausnahmefällen vorliegen ³⁾. Solche Ausnahmefälle werden in späteren Kapiteln eine zentrale Bedeutung erhalten.

Annahme $\beta_{(s)}$: $\text{plim}(\frac{1}{T} \underline{v}' \underline{v}) = \sigma^2(v)$ und $E(\underline{v} \underline{v}' / X) = \sigma^2(v) \cdot I$

Im Gegensatz zum Modell mit fixen Regressoren hat eine Verletzung vor allem des zweiten Teils dieser Annahme (keine Autokorrelation der Residuen) nicht nur Einfluß auf die Effizienz des OLS-Schätzers, sondern kann auch zum Verlust der Konsistenz führen. Sind nämlich z. B. unter den stochastischen Regressoren verzögerte endogene Variablen und sind gleichzeitig die v_t autokorreliert, so wird der OLS-Schätzer für $\hat{\gamma}$ inkonsistent ⁴⁾. Ebenso wie für das Modell fixer Regressoren ⁵⁾ existieren statistische Tests auf Autokorrelation; als Beispiel sei der

1) De-Min Wu (1973)

2) Eine Instrumentalvariable (IV) z ist definiert durch $\text{plim}(\frac{1}{T} \underline{z}' \underline{v}) = 0$; $\text{plim}(\frac{1}{T} \underline{z}' \underline{X}_{(s)}) = \underline{Z}_{zx}$ existiert und ist nicht singulär; $\text{plim}(\frac{1}{T} \underline{z}' \underline{z}) = \underline{Z}_{zz(s)}$ existiert

3) Außerdem ist dann ein konsistenter IV-Ansatz möglich; vgl. Kapitel 3

4) Vgl. z. B. Malinvaud (1970) S. 558 ff.

5) Eine Übersicht liefert z. B. Maddala (1977) S. 285 ff.

von Durbin ¹⁾ vorgeschlagene Test auf Autokorrelation genannt, der auch für Modelle mit verzögerten endogenen Variablen als Regressoren geeignet ist.

Annahme 4_(s): $\text{plim } (X'X) = \Sigma_{XX}$ existiert und ist nichtsingulär. Formal zum Nachweis der Konsistenz des OLS-Schätzers nötig ²⁾, ist diese Annahme bei der jeweils für die Schätzung vorliegenden Realisation der Zufallsvariablen $(X'X)$ nicht nachprüfbar. Wenn die zur Schätzung vorliegende, aus einer endlichen Anzahl von Messungen entstandene Matrix $X'X$ singulär ist, muß die OLS-Schätzung mit Hilfe der Pseudoinversen (Penroseinversen) von $X'X$ durchgeführt werden ³⁾, die allerdings keine eindeutige Bestimmung des Schätzvektors $\hat{\underline{\gamma}}$ erlaubt. In der ökonometrischen Schätzpraxis wird die unbekannte Matrix Σ_{XX} durch $X'X$ ersetzt, ebenso Σ_{XX}^{-1} durch $(X'X)^{-1}$. Dies gilt sowohl für die OLS-Schätzung selbst wie auch für die auf Σ_{XX}^{-1} aufbauenden Berechnungen wie z. B. $\sigma^2(\hat{\underline{\gamma}})$ ⁴⁾.

Annahme 5_(s): Die Randverteilung von $\underline{v}$ ist $(0; \sigma^2(v) \cdot I)$ -normal. Unter dieser Annahme gelten alle Tests und Konfidenzgrenzen, die auf den Verteilungen der Prüfmaße χ^2 , F und t beruhen, auch für den Fall stochastischer Regressoren ⁵⁾. Die von Maddala ⁶⁾ für den Fall fixer Regressoren angegebenen Tests auf Normalverteilung der Residuen sind nichtparametrische Tests und können in dem vorliegenden Fall ebenfalls verwendet werden.

1) Durbin (1970)

2) Vgl. z. B. Schönfeld (1971) Band II S. 24

3) Vgl. Schönfeld (1971) Band II S. 17 und S. 293 f.

4) Vgl. z. B. Maddala (1977) S. 152

5) Vgl. z. B. Schönfeld (1971) S. 22 f.

6) Maddala (1977) S. 306 f.

Die Ergebnisse dieses Kapitels lassen sich damit wie folgt zusammenfassen: Solange eine geschätzte Eingleichungsregression ein kleines Bestimmtheitsmaß aufweist, für das man den Meßfehler bei dem Regressanden nicht oder nur in geringem Maße verantwortlich machen kann, müssen die bei der OLS-Schätzung zu treffenden Annahmen sorgfältig geprüft werden im Sinne einer Plausibilitätsprüfung. Dabei muß man jedoch, vor allem bei dem für ökonometrische Untersuchungen typischen Datenmaterial, im Zweifelsfall wohl die Annahmen als unzutreffend verwerfen und bei der Interpretation der Ergebnisse entsprechend zurückhaltend argumentieren ¹⁾. In diesem Zusammenhang prangert Liu ²⁾ ein in der Schätzpraxis oft geübtes Verfahren an, das bei im Widerspruch zu einer (z. B. volkswirtschaftlichen) Hypothese stehendem Schätzergebnis versucht, durch Ausschluß von Regressoren aus dem Schätzansatz Übereinstimmung zwischen Schätzung und Hypothese zu erreichen: Ein wegen falscher (z. B. unvollständiger) Spezifikation verzerrter Schätzansatz wird nicht dadurch "richtiger", daß man durch den Ausschluß weiterer Regressoren für eine Kompensation der ursprünglichen Verzerrung sorgt. Umgekehrt darf natürlich nicht daraus geschlossen werden, daß die Schätzkoeffizienten eines mit möglichst vielen Regressoren spezifizierten Modells die "richtigeren" sind, auch wenn diese Schätzung nicht mit den hypothetischen Größen oder Vorzeichen übereinstimmt. Denn ein solches Modell kann zwar "more nearly correct specified" ³⁾ sein, aber es liefert eben immer noch verzerrte Schätzungen, da es nicht korrekt ist. Bevor man

1) Maddala (1977) S. 303 meint sogar generell: "With economic data the assumptions of zero means of the errors and of zero covariances between the errors and the systematic parts ... may not be valid assumptions to make".

2) Ta-Chung Liu (1960)

3) Ta-Chung Liu (1960) S. 859 Fußnote 8 Abschnitt 5.)

also eine Hypothese aufgrund eines Schätzansatzes verwirft, muß man unbedingt prüfen, ob eine mögliche Korrelation der in v subsumierten ausgeschlossenen Variablen eine zu diesem Widerspruch führende Verzerrung hervorgerufen haben kann. Ansätze hierzu liefern z. B. Wold & Faxer ¹⁾ bzw. Fisher ²⁾. Auf keinen Fall ist es zulässig, sich durch irgendeinen Schätzansatz von einem Datensatz eine Hypothese "formulieren" zu lassen und diese "Hypothese" dann durch denselben Datensatz oder, durch eine nur geringfügige Veränderung der Stichprobe, ähnlichen Datensatz zu "verifizieren".

In den folgenden Kapiteln wird daher immer davon ausgegangen, daß für den Schätzansatz eine zu testende Hypothese vorliegt, die eine bestimmte Spezifikation des Modells verlangt. Untersucht werden dann Modelle, die im Sinne Liu's "more nearly correct", aber eben nur beinahe korrekt sind. Unter zusätzlichen Annahmen soll dann versucht werden, diesen Fehler ganz oder zumindest teilweise zu beheben.

1) Wold & Faxer (1957)

2) F.M. Fisher (1961)

§2. Verletzung der Modellannahmen des OLS-Schätzers durch
=====
falsche Wahl der Regressorenmenge
=====

In Kapitel 1 wurde behauptet, daß bestimmte Verletzungen der für eine OLS-Schätzung notwendigen Modellannahmen zu verzerrten und inkonsistenten Schätzungen führen können. Dabei wurde vor allem die zentrale Bedeutung der Annahme 2 (bzw. Annahme 2_(s)) $E(v_t/x_{it}) = 0$ betont. Die folgenden Kapitel werden ebenso wie das vorliegende Spezialfälle solcher Verletzungen behandeln, die der Verfasser wesentlich für die Schätzpraxis hält.

Ein solcher Sonderfall ist eine vom "wahren" Modell abweichende Wahl der Menge der Regressoren ¹⁾. Ob und in welcher Form eine solche Fehlspezifikation zu einer Verletzung der Annahme 2 führen kann, wird anhand dreier Fälle untersucht:

- a) "Unterspezifizierte" Regressionsansätze, d. h. nur mit einer Teilmenge der "wahren" Regressoren spezifizierte OLS-Schätzansätze
- b) "Überspezifizierte" Regressionsansätze, d. h. mit einer Menge von Regressoren spezifizierte OLS-Schätzansätze, die die "wahren" Regressoren als Teilmenge beinhaltet
- c) Mischformen der Fehlspezifikationen nach a) und b), d. h. es werden sowohl "wahre" Regressoren im Schätzansatz vernachlässigt wie auch über das "wahre" Modell hinausgehende Regressoren in den OLS-Schätzansatz einbezogen: Die Schnittmenge zwischen den "wahren" und den in den Schätzansatz einbezogenen Regressoren umfaßt weder die Menge der "wahren" noch die Menge der für den Schätzansatz benutzten Regressoren.

1) Im Gegensatz zu der im nächsten Kapitel behandelten Wahl des falschen Regressorentyps ("Fehler in den Regressoren") wird hier also die falsche Regressorenmenge gewählt.

2.1. Definition der Fehlertypen

Gegeben sei das "wahre" Modell $\underline{y} = \underline{X}\underline{\gamma} + \underline{v}$,

wobei die Dimensionen der Matrizen $\underline{y}_{T;1}$; $\underline{X}_{T;K}$; $\underline{\gamma}_{K;1}$

und $\underline{v}_{T;1}$ sind. Die Modellannahmen 1, 2 und 3 seien erfüllt.

Dann ist der OLS-Schätzer $\hat{\underline{\gamma}} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{y}$ ein Schätzer für

$\underline{\gamma}$ mit den BLUE-Eigenschaften: Unter allen linearen, erwartungs-
treuen Schätzern für $\underline{\gamma}$ hat jedes $\hat{\underline{\gamma}}_i$ die kleinste Varianz

($i=1, \dots, K$) ^①, nämlich $\text{var}(\hat{\underline{\gamma}}) = \sigma^2(\underline{X}'\underline{X})^{-1}$. Dabei läßt
sich σ^2 erwartungstreu schätzen nach $E(\hat{\sigma}^2) = E(s^2) = E(\underline{v}'\underline{v}) =$
 $= (T - K)\sigma^2$ mit $\hat{\underline{v}} = \underline{y} - \underline{X}\hat{\underline{\gamma}}$ ^②.

Versteht man unter Spezifikation lediglich die Auswahl der
im Schätzansatz einbezogenen exogenen Variablen, so lassen
sich vier Spezifikationsfälle unterscheiden:

		Fälschliches Einbeziehen von Var.	
		nein	ja
Fälschliches Ausschließen von Variablen	nein	BLUE des Modells	Fehlertyp I
	ja	Fehlertyp II	Fehlertyp III

Die Eigenschaften der Fehlertypen I, II und III werden im
folgenden detailliert dargestellt, wobei bezüglich der
Nomenklatur gelten soll:

$\underline{X} = (\underline{X}_1; \underline{X}_2)$ Die Matrix $\underline{X}$ setzt sich aus den beiden Teil-
matrizen $\underline{X}_1$ und $\underline{X}_2$ zusammen

$\underline{X}_1(T; K_1)$ ist die Teilmatrix von $\underline{X}$ mit $K_1 < K$, die in den
Schätzansatz einbezogen wird

$\underline{X}_2(T; K_2)$ ist die Teilmatrix von $\underline{X}$ mit $K_2 = K - K_1 < K$,
die bei den Fehlertypen II und III aus dem
Schätzansatz ausgeschlossen wird

$\underline{X}_3(T; K_3)$ ist die Matrix von x -Werten, die bei den Fehler-
typen I und III in den Schätzansatz einbezogen
wird mit $K_3 > 0$; $\underline{X}_3$ ist keine Teilmatrix von $\underline{X}$

"postuliert" wird eine Spezifikation dann genannt, wenn sie
sich von der des "wahren" Modells unterscheidet

① Vergl. z. B. SCHNEEWEIS (1974) S. 73

② Vergl. z. B. SCHNEEWEIS (1974) S. 109 f.

2.1.1. Fehlspezifikation nach Fehlertyp I

Das postulierte Modell laute

$$\underline{y} = (X; X_3) (\underline{\gamma}; \underline{\gamma}_3)' + \underline{v}^* \quad ; \text{ dies führt zu dem OLS-Ansatz}$$

$$(\hat{\underline{\gamma}}; \hat{\underline{\gamma}}_3) = [(X; X_3)'(X; X_3)]^{-1} (X; X_3)' \underline{y} \quad \text{oder ausführlicher}$$

$$\begin{pmatrix} \hat{\underline{\gamma}} \\ \hat{\underline{\gamma}}_3 \end{pmatrix} = \begin{pmatrix} X'X & X'X_3 \\ X_3'X & X_3'X_3 \end{pmatrix}^{-1} \begin{pmatrix} X'y \\ X_3'y \end{pmatrix}$$

Unter Anwendung der Regel für das Invertieren von partitionierten Matrizen ^① ist

$$\begin{pmatrix} X'X & X'X_3 \\ X_3'X & X_3'X_3 \end{pmatrix}^{-1} = \begin{bmatrix} (X'X)^{-1} + (X'X)^{-1}(X'X_3)H^{-1}(X_3'X)(X'X)^{-1} & -(X'X)^{-1}(X'X_3)H^{-1} \\ -H^{-1}(X_3'X)(X'X)^{-1} & H^{-1} \end{bmatrix}$$

$$\text{mit } H^{-1} = [(X_3'X_3) - (X_3'X)(X'X)^{-1}(X'X_3)]^{-1}$$

Berücksichtigt man hierbei, daß $E(\underline{y}) = X\underline{\gamma}$ ist, erhält man

$$E\left(\begin{pmatrix} \hat{\underline{\gamma}} \\ \hat{\underline{\gamma}}_3 \end{pmatrix}\right) = [(X; X_3)'(X; X_3)]^{-1} \begin{pmatrix} X'X \\ X_3'X \end{pmatrix} \underline{\gamma} =$$

$$= \begin{pmatrix} \underline{\gamma} + (X'X)^{-1}(X'X_3)H^{-1}(X_3'X)\underline{\gamma} - (X'X)^{-1}(X'X_3)H^{-1}(X_3'X)\underline{\gamma} \\ -H^{-1}(X_3'X)(X'X)^{-1}(X'X)\underline{\gamma} + H^{-1}(X_3'X)\underline{\gamma} \end{pmatrix} = \begin{pmatrix} \underline{\gamma} \\ \emptyset \end{pmatrix}$$

Wie man sieht, führt ein durch das "wahre" Modell nicht gerechtfertigtes Einführen zusätzlicher Regressoren nicht zu einer Verzerrung der geschätzten Regressionskoeffizienten. Dies bedeutet nun allerdings nicht, daß eine Fehlspezifikation vom Typ I ohne (negative!) Folgen bleibt. Dies wird durch einen Effizienzvergleich sofort deutlich:

Sei R_X die Korrelationsmatrix der Exogenen des "wahren" Modells

$R_{X;X_3}$ die Korrelationsmatrix aller Exogenen des postulierten Modells

Dann ist die relative Effizienz des Schätzers $\hat{\underline{\gamma}}_I$, verglichen mit $\hat{\underline{\gamma}}$

$$\text{Rel. Eff.}(\hat{\underline{\gamma}}_I) = \frac{\text{diag}(\text{var}(\hat{\underline{\gamma}}))}{\text{diag}(\text{var}(\hat{\underline{\gamma}}_I))} = \frac{\text{diag}(R_X)^{-1}}{\text{diag}(R_{X;X_3})^{-1}} \quad ,$$

wobei diag die Diagonalelemente der betreffenden Matrix kennzeichnet und die Division elementweise vorgenommen wird ^②

1) Vergl. z. B. SCHNEEWEIS (1974) S. 335

2) Zur Definition der relativen Effizienz für unverzerrte Schätzer vergl. z. B. WONNACOTT & WONNACOTT (1970) S. 43

Daß $\text{diag}(R_X)^{-1} < \text{diag}(R_{X;X_3})^{-1}$ ist ^①, sieht man an der

partitionierten Matrix:

$$\text{diag}(X'X)^{-1} < \text{diag}((X'X)^{-1} + (X'X)^{-1}(X'X_3)H^{-1}(X_3'X)(X'X)^{-1})$$

Entsprechend ist sofort zu sehen, daß gilt ^②

$$\text{Mean square error}(\hat{\gamma}_I) = \text{MSE}(\hat{\gamma}_I) \geq \text{MSE}(\hat{\gamma})$$

Jedoch wird $\hat{\sigma}^2$ durch $\hat{\sigma}^2 = s^2$ unverzerrt geschätzt ^③:

$$\hat{\underline{v}} = \underline{y} - (X;X_3)(\hat{\underline{\gamma}}_I; \hat{\underline{\gamma}}_3) = \{I - (X;X_3) [(X;X_3)'(X;X_3)]^{-1}(X;X_3)'\} \underline{y}$$

= $\underline{M}\underline{y}$. Da aber $\underline{y} = \underline{X}\underline{\gamma} + \underline{v}$, ist $\hat{\underline{v}} = \underline{M}\underline{X}\underline{\gamma} + \underline{M}\underline{v}$ und wegen

$$\underline{M}\underline{X} = \underline{0} \text{ ist dann } \hat{\underline{v}} = \underline{M}\underline{v}$$

Damit ist $\hat{\sigma}^2 = \hat{\underline{v}}'\hat{\underline{v}}/(T - K - K_3)$. Also gilt

$$E(\hat{\sigma}^2) = \frac{1}{(T - K - K_3)} E(\underline{v}'\underline{M}\underline{v}) = \frac{(T - K - K_3)}{(T - K - K_3)} \sigma^2 = \sigma^2$$

2.1.2. Fehlspezifikation nach Fehlertyp II

Schreibt man das "wahre" Modell $E(\underline{y}) = X_1\underline{\gamma}_1 + X_2\underline{\gamma}_2$, wobei X_2 mindestens den Rang 1 haben soll und weder $\underline{\gamma}_1$

noch $\underline{\gamma}_2$ Nullelemente enthält, so führt ein nach

$E(\underline{y}) = X_1\underline{\gamma}_1^{\text{II}}$ postuliertes Modell zu dem Schätzansatz

$$\hat{\underline{\gamma}}_1^{\text{II}} = (X_1'X_1)^{-1}X_1'\underline{y}$$

Daß dies ein verzerrter (und selbstverständlich auch inkonsistenter) Schätzer für $\underline{\gamma}_1$ ist, sieht man sofort nach

$$E(\hat{\underline{\gamma}}_1^{\text{II}}) = \underline{\gamma}_1 + (X_1'X_1)^{-1}X_1'X_2\underline{\gamma}_2$$

, wenn nicht alle Vektoren von X_2 orthogonal zu X_1 sind.

Ein Effizienzvergleich ist wegen der Verzerrtheit von $\hat{\underline{\gamma}}_1^{\text{II}}$ nicht sinnvoll; da aber z. B. schrittweise Spezifikationsverfahren auf dem MSE aufbauen ^④, soll für diesen eine Gegenüberstellung gegeben werden ^⑤:

① Da angenommen wird, daß $R_{X;X_3}$ den Rang $(K + K_3)$ hat, kann man den Fall $\text{diag}(R_X)^{-1} = \text{diag}(R_{X;X_3})^{-1}$ übergehen.

② Vergl. MADDALA (1977) S. 157

③ Vergl. MADDALA (1977) S. 157

④ Vergl. z. B. SCHNEEWEIS (1974) S. 154

⑤ Vergl. LEAMER (1978) S. 131

$$\text{MSE}(\hat{\underline{\gamma}}_1; \hat{\underline{\gamma}}_2) = \sigma^2 \begin{pmatrix} \underline{X}_1' \underline{X}_1 & \underline{X}_1' \underline{X}_2 \\ \underline{X}_2' \underline{X}_1 & \underline{X}_2' \underline{X}_2 \end{pmatrix}^{-1} \quad \text{und}$$

$$\text{MSE}(\hat{\underline{\gamma}}_1^{\text{II}}) = \sigma^2 \begin{pmatrix} (\underline{X}_1' \underline{X}_1)^{-1} & \emptyset \\ \emptyset & \emptyset \end{pmatrix} + \underline{\gamma}_2^2 \begin{pmatrix} (\underline{X}_1' \underline{X}_1)^{-2} (\underline{X}_1' \underline{X}_2)^2 & -(\underline{X}_1' \underline{X}_1)^{-1} (\underline{X}_1' \underline{X}_2) \\ -(\underline{X}_1' \underline{X}_1)^{-1} (\underline{X}_1' \underline{X}_2) & 1 \end{pmatrix}$$

So würde der Schätzer $\hat{\underline{\gamma}}_1^{\text{II}}$ dem Schätzer $\hat{\underline{\gamma}}_1$ nach dem MSE-Kriterium vorgezogen werden, wenn

$$\sigma^2 (\underline{X}_1' \underline{X}_1)^{-1} + \hat{\underline{\gamma}}_2^2 (\underline{X}_1' \underline{X}_1)^{-2} (\underline{X}_1' \underline{X}_2)^2 < \sigma^2 [(\underline{X}_1' \underline{X}_1 - \underline{X}_1' \underline{X}_2 (\underline{X}_2' \underline{X}_2)^{-1} \underline{X}_2' \underline{X}_1)]^{-1}.$$

Wie vor allem Rao ^① zeigt, sind die aufgrund des postulierten Modells gewonnenen Schätzungen nicht nur verzerrt, sondern werden als hochsignifikant getestet wegen der reduzierten Varianz-Kovarianzmatrix.

2.1.3. Fehlspezifikation nach Fehlertyp III

Unterteilt man das "wahre" Modell wie im vorhergehenden Abschnitt in zwei Teile, so läßt sich das postulierte Modell schreiben

$$E(\underline{y}) = \underline{X}_1 \underline{\gamma}_1^{\text{III}} + \underline{X}_3 \underline{\gamma}_3^{\text{III}}. \quad \text{Fügt man dem "wahren" Modell}$$

für den Vergleich noch die Matrix $\underline{X}_3$ und den aus Nullelementen bestehenden Koeffizientenvektor $\underline{\gamma}_3$ hinzu, so erhält man entsprechend dem Vorgehen in Abschnitt 2.1.2. den verzerrten Schätzer

$$E(\hat{\underline{\gamma}}_1^{\text{III}}; \hat{\underline{\gamma}}_3^{\text{III}}) = (\underline{\gamma}_1; \emptyset) + [(\underline{X}_1; \underline{X}_3)' (\underline{X}_1; \underline{X}_3)]^{-1} (\underline{X}_1; \underline{X}_3)' \underline{X}_2 \underline{\gamma}_2.$$

Auch hier tritt, wie beim Fehlertyp II, in dem Bias die "Hilfsregression" der nicht in das Modell einbezogenen Variablen $\underline{X}_2$ auf die im Modell postulierten auf, hier also $\underline{X}_1$ und $\underline{X}_3$.

① RAO (1971)

§3. Verletzung der Modellannahmen des OLS-Schätzers durch
=====

Verwendung eines unzulässigen Regressortyps
=====

Wie schon in Kapitel 2 soll auch in diesem Kapitel eine Ursache für die Verletzung der Forderung $E(v_t/x_{1t}) = 0$ untersucht werden. Eine mögliche Ursache für eine Verletzung dieser Annahme 2 liegt in dem sogenannten "Fehler-in-den-regressoren"-Modell ¹⁾, von dem z. B. Schönfeld meint ²⁾, daß es im Vergleich zu dem "Fehler-in-den-Gleichungen"-Modell ³⁾ in der Praxis viel häufiger benutzt werden müßte. Derselbe liefert auch die Begründung für das seltene Antreffen eines EIV-Ansatzes ¹⁾: Er bereitet erhebliche Schwierigkeiten bei der Parameterschätzung. Diese Schwierigkeiten liegen sowohl in der Entwicklung eines geeigneten Schätzansatzes wie auch - und vor allem hier tauchen in der Schätzpraxis die Probleme auf - in der Plausibilität der für diese Schätzansätze nötigen Modellannahmen.

Die folgenden Abschnitte sollen die Auswirkungen eines OLS-Ansatzes auf ein EIV-Modell zeigen; typische und auch versteckte Formen fehlerbehafteter Regressoren werden vorgeführt und interpretiert. Nach der Diskussion der Frage, ob und wann die Verwendung fehlerbehafteter Regressoren sinnvoll sein kann, werden Lösungsansätze für EIV-Modelle behandelt.

Im Vorgriff auf spätere Kapitel soll jetzt schon betont werden ⁴⁾, daß zwar EIV-Modelle und interdependente Strukturgleichungen sowohl die Verletzung der Annahme 2 wie auch bestimmte Schätzansätze ⁵⁾ gemeinsam haben, darüber hinaus jedoch keine formalen oder interpretatorischen Übereinstimmungen existieren.

1) In der angelsächsischen Literatur findet man hierfür oft die Abkürzung **EIV**-(Errors In the Variables)-model, die deshalb auch im folgenden verwendet wird

2) Schönfeld (1971) Band II S. 106

3) Im angelsächsischen "error shock model"; entspricht Fehler-b) des Kapitels 1 S. 3

4) Vor allem Kapitel 7

5) Nämlich die Instrumentalvariablen-Schätzung

3.1. Einregressormodell mit fehlerbehaftetem Regressor

Gegeben sei das "wahre" Modell ¹⁾

$$\underline{y} = \underline{x} \gamma + \underline{v}$$

für das die Annahmen 2, 3 und 4 erfüllt seien. Gemessen sind die Werte des Vektors $\underline{y}$ und, anstelle des Regressors $\underline{x}$, eine mit Fehlern behaftete Variable $\underline{x}^+ = \underline{x} + \underline{v}^+$. Trifft man, wie für das "Fehler-in-dem-Regressor"-Modell üblich, die Annahmen

$$G(\underline{v}, \underline{v}^+) = G(\underline{x}, \underline{v}^+) = E(v_t^+) = 0$$

so sieht man ²⁾, daß die Verwendung des OLS-Schätzers unter Verwendung von $\underline{x}^+$ zu inkonsistenten Schätzungen für γ führt:

$$\text{plim}(\hat{\gamma}^+) = \frac{\text{cov}(\underline{x}^+, \underline{y})}{\text{var}(\underline{x}^+)} = \frac{\text{cov}(\underline{x}, \underline{y})}{\text{var}(\underline{x}) + \text{var}(\underline{v}^+)} = \gamma \frac{1}{1 + \frac{\text{var}(\underline{v}^+)}{\text{var}(\underline{x})}} = \gamma \frac{\text{var}(\underline{x})}{\text{var}(\underline{x}^+)}$$

Der OLS-Schätzer ist also gegen 0 verzerrt, γ wird also unterschätzt. Die Varianz des Residuums v jedoch wird überschätzt nach

$$\text{plim}(\hat{\sigma}^2) = \text{plim}\left(\frac{1}{n-2} \sum_t (y_t - \hat{\gamma}^+ x_t^+)^2\right) = \text{var}(\underline{y}) + \text{var}(\underline{v}) - \gamma^2 \frac{\text{var}(\underline{x})}{\text{var}(\underline{x}^+)} \cdot \text{var}(\underline{x}^+)$$

$$= \text{var}(\underline{v}) + \left(1 - \frac{\text{var}(\underline{x})}{\text{var}(\underline{x}^+)}\right) \gamma^2 \text{var}(\underline{x}) = \text{var}(\underline{v}) + \frac{\text{var}(\underline{v}^+)}{\text{var}(\underline{x}^+)} \gamma^2 \text{var}(\underline{x})$$

Während die Verzerrung von $\hat{\gamma}$ also lediglich von dem Varianzanteil des Fehlers $\underline{v}^+$ an der Variablen $\underline{x}^+$ bestimmt wird, wird bei der Verzerrung von $\hat{\sigma}^2$ dieses Varianzverhältnis noch mit der "erklärten" Varianz von $\underline{y}$ gewichtet, also der Varianz von $\underline{y}$.

3.2. Mehrregressormodell mit einem fehlerbehafteten Regressor ³⁾

In diesem Modell seien nun neben anderen ein fehlerbehafteter Regressor, z. B. der erste (o.B.d.A.):

$$\underline{y} = \underline{x}_1 \gamma_1 + \sum_{i=2}^K \underline{x}_i \gamma_i + \underline{v} = \underline{X} \underline{\gamma} + \underline{v}$$

1) Zu beachten ist, daß wiederum alle Vektoren als Abweichungen von ihrem arithmetischen Mittel dargestellt sind

2) Vgl. z. B. Maddala (1977) S. 292 f.

3) Vgl. z. B. M. D. Levi (1973, 1977) und Griliches (1974)

Gemessen werden, neben den Vektoren $\underline{y}$ und $\underline{x}_i$ ($i=2, \dots, K$), statt des Vektors $\underline{x}_1$ der Vektor $\underline{x}_1^+ = \underline{x}_1 + \underline{v}^+$, wobei die Unkorreliertheit der Fehler $\underline{v}$ und $\underline{v}^+$ untereinander und zu allen Regressoren $\underline{x}_i$ ($i=1, \dots, K$) wie in Abschnitt 3.1. gelten soll. Die Matrix $(\underline{x}_1^+, \underline{x}_2, \dots, \underline{x}_K)$ werde mit $\underline{X}^+$ bezeichnet. Dann führt der OLS-Ansatz zu einem $\hat{\underline{\gamma}}^+$ mit den Eigenschaften:

$$\text{plim}(\hat{\underline{\gamma}}^+) = \text{plim}(\underline{X}^+{}' \underline{X}^+)^{-1} \underline{X}^+{}' \underline{y} = \text{plim}\left(\frac{\underline{X}^+{}' \underline{X}^+}{T}\right)^{-1} \text{plim}\left(\frac{\underline{X}^+{}' \underline{y}}{T}\right) =$$

$$= (\underline{\Sigma} + \underline{\Omega})^{-1} \underline{\Sigma} \underline{\gamma}$$

wenn die $\underline{x}_i$ ($i=1, \dots, K$) und $\underline{v}^+$ die endlichen Varianz-Kovarianz-Matrizen $\underline{\Sigma}$ und $\underline{\Omega}$ besitzen.

Levi ¹⁾ zeigt die Abhängigkeit von $\text{plim}(\hat{\underline{\gamma}}^+)$ als Funktion einer Veränderung von $\sigma^2(\underline{v}^+)$ nach

$$\frac{d(\text{plim} \hat{\underline{\gamma}}^+)}{d(\sigma^2(\underline{v}^+))} = -(\underline{\Sigma} + \underline{\Omega})^{-1} \begin{pmatrix} \text{plim}(\hat{\underline{\gamma}}^+) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (\text{I})$$

Da $(\underline{\Sigma} + \underline{\Omega})^{-1}$ positiv definit ist, ist das Element in der linken oberen Ecke positiv, das Vorzeichen des ersten Gliedes in (I) ist also dem von $\text{plim}(\hat{\underline{\gamma}}^+)$ entgegengesetzt. Der zum fehlerbehafteten Regressor gehörende Koeffizient γ_1 ist also auch hier gegen Null verzerrt. Die Koeffizienten der fehlerfrei gemessenen Regressoren sind ebenfalls **verzerrt**, die Richtung der Verzerrung ergibt sich aus folgender Überlegung:

Die Richtungen der Verzerrungen von γ_2 bis γ_K hängen nur von den Vorzeichen der ersten Spalte von $(\underline{\Sigma} + \underline{\Omega})^{-1}$ ab. Da aber $\underline{\Omega}$ bis auf die linke obere Ecke lediglich Nullen aufweist, erhält man bei Anwendung der Kofaktormethode beim Invertieren von $(\underline{\Sigma} + \underline{\Omega})$ die Elemente der ersten Spalte der Inversen zu $\underline{\Sigma}_{i1} / \text{Det}|\underline{\Sigma} + \underline{\Omega}|$ mit $\underline{\Sigma}_{i1}$ als dem $(1, i)$ -ten Kofaktor von $\underline{\Sigma}$, der für alle i nur Nullelemente von $\underline{\Omega}$ beinhaltet.

1) Levi (1973)

Damit hängen die Vorzeichen der ersten Spalte von $(\sum + \Omega)^{-1}$ nur von $\sum$ ab, die Richtung der Verzerrung der $\hat{\gamma}_i$ ($i=2, \dots, K$) hängt also von den Vorzeichen der γ_i ($i=2, \dots, K$) und dem Vorzeichen von γ_1 ab.

Dies soll an einem kleinen Beispiel verdeutlicht werden ¹⁾:

Die wahre Beziehung laute $y = x_1 \gamma_1 + x_2 \gamma_2 + x_3 \gamma_3 + v$, die Variable x_1^+ sei der fehlerbehaftete Regressor nach

$x_1^+ = x_1 + v^+$. Die Modellannahmen seien alle erfüllt. Dann

findet man für $\text{plim}(\hat{\gamma}^+)$ mit $\text{DET} = \text{Determinante von } (X'X)$,

$\text{DET}^+ = \text{Determinante von } (X^{+'}X^+)$

$$\begin{aligned} \text{plim}(\hat{\gamma}^+) &= \begin{bmatrix} \sigma_{11}^2 + \sigma_{v^+}^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}^{-1} X^{+'} y \\ &= \frac{1}{\text{DET} + \sigma_{v^+}^2 \sigma_{22} \sigma_{33} - \sigma_{v^+}^2 \sigma_{23} \sigma_{32}} \left[\text{DET} \cdot (X'X)^{-1} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \sigma_{v^+}^2 \sigma_{22} & -\sigma_{v^+}^2 \sigma_{23} \\ 0 & -\sigma_{v^+}^2 \sigma_{23} & \sigma_{v^+}^2 \sigma_{33} \end{bmatrix} X' y \right] \\ &= \frac{\text{DET}}{\text{DET}^+} \gamma + \frac{\sigma_{v^+}^2}{\text{DET}^+} \begin{bmatrix} 0 & 0 & 0 \\ \sigma_{12} \sigma_{22} - \sigma_{23} \sigma_{13} & \sigma_{22} \sigma_{23} - \sigma_{23} \sigma_{23} & \sigma_{22} \sigma_{23} \\ -\sigma_{22} \sigma_{12} + \sigma_{22}^2 \sigma_{13} & 0 & -\sigma_{23} \sigma_{23} + \sigma_{22}^2 \sigma_{33} \end{bmatrix} \gamma \\ &= \frac{1}{\text{DET}^+} \begin{bmatrix} \text{DET} + \sigma_{v^+}^2 & 0 & 0 \\ 0 & \sim & 0 \\ 0 & 0 & \sim \end{bmatrix} \end{aligned}$$

Damit aber erhält man für

$$\begin{aligned} \text{plim}(\hat{\gamma}_1) &= \frac{\text{DET}}{\text{DET}^+} \gamma_1 \\ \text{plim}(\hat{\gamma}_2) &= \gamma_2 + \frac{\sigma_{v^+}^2 (\sigma_{12} \sigma_{22} - \sigma_{23} \sigma_{13})}{\text{DET}^+} \gamma_1 \\ \text{plim}(\hat{\gamma}_3) &= \gamma_3 + \frac{\sigma_{v^+}^2 (\sigma_{13} \sigma_{22} - \sigma_{23} \sigma_{12})}{\text{DET}^+} \gamma_1 \end{aligned}$$

$$\begin{aligned} \text{Für } \text{DET}^+ \text{ gilt aber } \text{DET}^+ &= \text{DET} + \sigma_{v^+}^2 \sigma_{22} \sigma_{33} - \sigma_{v^+}^2 \sigma_{23} \sigma_{23} = \\ &= \text{DET} + \sigma_{v^+}^2 \sigma_{22} \sigma_{33} (1 - r_{23}^2) \geq \text{DET} \end{aligned}$$

1) Ein Zahlenbeispiel hierzu ist im Anhang S. A.I.

An dieser Stelle soll aber darauf hingewiesen werden, daß die Verzerrung ausschließlich die Schätzung der Regressionskoeffizienten berührt. Ist man lediglich an einem Prognoseansatz interessiert, so liefert der OLS-Schätzer von $\underline{y}$ auf X^+ den optimalen Prognoseansatz ¹⁾.

3.3. Mehrregressormodell mit mehreren fehlerbehafteten Regressoren

Im Gegensatz zu den beiden bisher behandelten Fällen, in denen zwar nicht die Größe, aber doch wenigstens die Richtung der Verzerrung bestimmt werden konnte, ist dies bei Modellen mit mehreren fehlerbehafteten Regressoren praktisch nicht mehr möglich. Näherungsformeln für den noch verhältnismäßig überschaubaren Fall eines Modells mit zwei fehlerbehafteten Regressoren findet man bei Theil ²⁾. Danach gilt näherungsweise für ein Modell, das nur aus 2 fehlerbehafteten Regressoren besteht und alle Modellannahmen nach 3.2. erfüllt

$$\text{plim}(\hat{\gamma}_1^+) \approx \gamma_1 - \frac{1}{1-\varrho^2} \left(\frac{\sigma_{v_1}^2}{\sigma_{x_1}^2} \gamma_1 - \frac{\sigma_{v_2}^2}{\sigma_{x_2}^2} \gamma_2 \right) \quad \text{und analog}$$

$$\text{plim}(\hat{\gamma}_2^+) \approx \gamma_2 - \frac{1}{1-\varrho^2} \left(\frac{\sigma_{v_2}^2}{\sigma_{x_2}^2} \gamma_2 - \frac{\sigma_{v_1}^2}{\sigma_{x_1}^2} \gamma_1 \right)$$

wobei ϱ hierbei der Korrelationskoeffizient zwischen $\underline{x}_1$ und $\underline{x}_2$ ist.

Mit diesem Ergebnis ist zweierlei zu sehen: Erstens ist wegen der Unmöglichkeit, die Inverse von $(X^+)'X^+$ für beliebig viele Regressoren so anschaulich wie in Abschnitt 3.2. aufzuspalten, eine genaue und allgemeine Angabe der Richtung der Verzerrung von $\hat{\underline{\gamma}}^+$ nicht mehr möglich. Zweitens wird deutlich, daß bei Anwesenheit von mehr als einem fehlerbehafteten Regressor die Ergebnisse einer OLS-Schätzung mit den zu schätzenden Koeffizienten

1) Vgl. z. B. Wonnacott & Wonnacott (1970) S. 165 oder Johnston (1972) S. 290

2) Theil (1961) S. 328 f.

weder in ihrer Größe noch in ihrem Vorzeichen übereinstimmen müssen: Eine Interpretation der so gewonnenen Schätzungen wird unmöglich, das Testen einer (z. B. volkswirtschaftlichen) Hypothese hat keine Aussagekraft mehr.

3.4. Die Bedeutung fehlerbehafteter Regressoren

Fehlerbehaftete Regressoren kann man in zwei Gruppen unterteilen, die unterschiedliche Gründe für das Auftreten des Fehlers aufweisen: Meßfehlerbehaftete Variablen ¹⁾ und "Proxy-Variablen". Diese beiden Gruppen sollen im Folgenden deutlich differenziert werden, da sie bei der Verwendung in einem Schätzansatz eine unterschiedliche Behandlung verlangen.

3.4.1. Proxy-Variablen

Unter "Proxy"- oder "Stellvertreter"-Variablen versteht man Variablen, die anstelle einer nicht beobachtbaren oder nicht meßbaren ²⁾ fiktiven Variablen in den Regressionsansatz einbezogen werden. Das klassische Beispiel ist die "permanent - income" -Hypothese von Friedman ³⁾, welche eine Abhängigkeit des 'normalen', d. h. vom Konsumenten als normal empfundenen Konsums von einem fiktiven "Dauereinkommen" postuliert. Dabei ist das "Dauereinkommen" nicht meßbar, sondern besteht nur in der (bewußten oder unbewußten) Vorstellung des Konsumenten und kann für einzelne Zeitpunkte von dem gemessenen Einkommen erheblich abweichen. Eine Proxy ist also eine Ersatzvariable für eine andere nicht verfügbare "Original"-Variable und soll dieser möglichst 'ähnlich' sein, d. h. die beiden Variablen sollen hoch korreliert sein. Dabei kann die Einführung der Proxy

-
- 1) In Anlehnung an das angelsächsische Schrifttum wird hier lediglich von Variablen gesprochen, obwohl hier grundsätzlich immer Regressoren angesprochen sind
 - 2) Im angelsächsischen Schrifttum findet man daher auch die Bezeichnungen "latent" bzw. "unobservable" variables.
 - 3) M. Friedman (1957)

zwei Gründe haben, die einzeln oder gemeinsam vorhanden sein können: Zum einen kann der Ökonometriker an den Regressionskoeffizienten der "Original"-Variablen selbst interessiert sein, zum anderen kann die Einführung der Proxy eine Verzerrung der restlichen Schätzkoeffizienten des Regressionsansatzes verringern, die ja sonst bei Ausschluß einer Variablen auftritt, wenn diese mit den restlichen Regressoren korreliert ist ¹⁾.

Vor allem den zweiten Punkt sprechen McCullum und Wickens ²⁾ an, wenn sie beweisen, daß die Einführung selbst einer 'schlechten' ³⁾ Proxy in einen einfachen OLS-Ansatz die Verzerrung vermindert. Das Ergebnis des Beweises von McCullum sei hier kurz dargestellt ⁴⁾:

Für das Modell $\underline{y} = \underline{X}\underline{\beta} + \underline{x}_1\beta_1 + \underline{v}$, in dem neben der abhängigen Variablen $\underline{y}$ die fehlerfreien Regressoren $\underline{x}_j$ ($j=2, \dots, K$) zur Matrix $\underline{X}$ zusammengefaßt sind und für dessen OLS-Schätzung statt des Regressors $\underline{x}_1$ nur eine Proxy $\underline{p} = \underline{x}_1 + \underline{v}^+$ vorliegt, sind bei Erfüllung der Modellannahmen aus 3.2. zu finden:

$\text{plim}(\hat{\underline{\beta}} - \underline{\beta}) = \beta_1(\underline{X}'\underline{X})^{-1}\underline{X}'\underline{x}_1$ bei Regression ohne $\underline{x}_1$ bzw. $\underline{p}$

$\text{plim}(\hat{\underline{\beta}} - \underline{\beta}) = \beta_1(\underline{X}'\underline{X})^{-1}\underline{X}'\underline{x}_1 \frac{\sigma_{v^+}^2}{\delta}$ bei Regression auf alle K

Regressoren und Verwendung von $\underline{p}$ statt $\underline{x}_1$. Hierbei ist δ definiert nach $\delta = \sigma_{v^+}^2 + \sigma_{x_1}^2 \left[1 - (\underline{x}_1'\underline{X}(\underline{X}'\underline{X})^{-1}\underline{X}'\underline{x}_1) / \sigma_{x_1}^2 \right]$

Da aber der Ausdruck in eckigen Klammern nichts anderes als das Bestimmtheitsmaß einer Regression von $\underline{x}_1$ auf $\underline{X}$ darstellt, ist für alle Fälle, in denen überhaupt eine Verzerrung auftritt, ⁵⁾ diese Verzerrung bei Verwendung von $\underline{p}$ kleiner als

1) Vgl. Kapitel 2.1.2

2) McCullum (1972) und Wickens (1972)

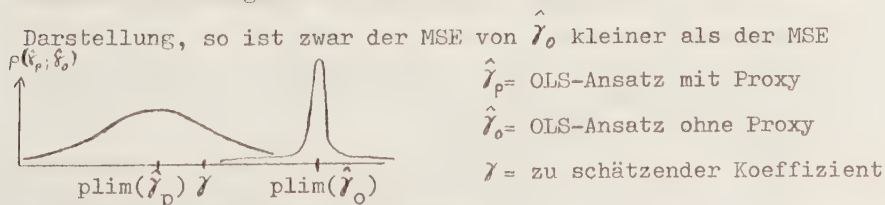
3) D. h. Proxy und "Original"-variable sind gering korreliert

4) Dies ist mit den Ergebnissen der Kapitel 2 und 3 leicht nachzuvollziehen

5) D. h., wenn $\underline{X}'\underline{x}_1 \neq 0$ ist

bei einer OLS-Schätzung ohne p.

Diese Feststellung wird für spätere Kapitel wichtig werden, ihre Bedeutung für die Schätzpraxis ist allerdings nicht unumstritten. Denn Aigner ¹⁾ zeigt, daß der 'mean square error' MSE der Regressionskoeffizienten $\hat{\gamma}_2$ bis $\hat{\gamma}_K$ bei einem OLS-Ansatz ohne Verwendung der Proxy-Variablen kleiner sein kann als bei dem Schätzansatz unter Einbeziehung der Proxy. Dieses Ergebnis soll hier nicht vorgeführt werden, da der Verfasser die Argumentation über den MSE als für die Schätzpraxis nicht relevant ablehnt. Dies hat zwei Gründe: Erstens kann bei einer durch Ausschluß eines Regressors fehlspezifizierten OLS-Schätzung das Vorzeichen der Schätzkoeffizienten umkippen, sodaß eine solche Schätzung nicht einmal zu einem Hypothesentest bezüglich der Vorzeichen taugt. Zweitens werden in der Schätzpraxis neben der Punktschätzung der Regressionskoeffizienten immer gleichzeitig Vertrauensintervalle für diese Punktschätzungen angegeben. Diese Vertrauensintervalle können aber selbstverständlich nicht unter Verwendung der ja unbekannten Verzerrung angegeben werden, sondern werden ja unter der Annahme eines "richtigen" Ansatzes aufgrund der bekannten Formeln für eine korrekt spezifizierte OLS-Schätzung angegeben: Die bekannte Varianzzerlegung ²⁾ für den MSE eines Schätzers $\hat{\gamma}$ nach $MSE = \sigma^2(\hat{\gamma}) + (\text{bias})^2$ kann ja nicht nachvollzogen werden. Betrachtet man z. B. folgende ³⁾



1) D. J. Aigner (1974)

2) Vgl. z. B. Wonnacott & Wonnacott (1970) S. 46

3) Ein numerisches Beispiel findet sich im Anhang S. A. II.

von $\hat{\gamma}_p$, aber von dem Ökonometriker wird die Hypothese $\gamma = \hat{\gamma}_0$ mit einer hohen Wahrscheinlichkeit abgelehnt. Damit aber ist die Regression als Instrument zur Falsifikation von Hypothesen aufgrund empirischer Daten unbrauchbar geworden.

3.4.2. Meßfehlerbehaftete Regressoren

Im Gegensatz zu einer Proxy stellt eine meßfehlerbehaftete Variable keine unvollkommene "Stellvertreter"-Variable dar. Hier ist, zumindest dem Namen nach, genau die Variable in den Schätzansatz einbezogen, deren Koeffizient man schätzen möchte. Zur Schätzung dieses Ansatzes stehen allerdings nur Messungen dieser Variablen zur Verfügung, die mit meßfehlerbehaftet sind. Dieser Fall tritt in der ökonometrischen Schätzpraxis nicht selten auf; in Modellen mit makroökonomischen Variablen kann man ihn wohl sogar als den Regelfall betrachten. Denn bei makroökonomischen Variablen kann sowohl ein Aggregationsfehler wie auch ein "Schätz"-Fehler auftreten, da ja nicht alle Größen unmittelbar gemessen werden ¹⁾. Formal kann man eine meßfehlerbehaftete Variable ebenso wie eine Proxy behandeln, Unterschiede treten erst bei der Plausibilitätsprüfung der für konsistente Schätzansätze im EIV-Modell nötigen Annahmen auf.

3.5. Bekannte Schätzansätze für das "Fehler-in-den-Regressoren"-Modell (EIV-Modell)

Für die Darstellung der folgenden Schätzansätze wird immer das Modell mit einem Regressor verwendet; für die Verallgemeinerung auf das multiple Modell wird lediglich auf die zugehörige Literatur verwiesen.

3.5.1. Der Instrumentalvariablenansatz

Gegeben sei das Modell $\underline{y} = \underline{x}\gamma + \underline{v}$, das die Annahmen des Kapitels 1 erfüllt. Neben $\underline{y}$ werde statt der Variablen $\underline{x}$

1) Vgl. Fußnote 2) auf S. 3

eine Proxy $\underline{x}^p = \underline{p} = \underline{x} + \underline{v}^+ = \underline{x} + \underline{v}^p$ gemessen mit den Eigenschaften $E(\underline{v}^p, \underline{x}) = 0$, $E(\underline{v}^p, \underline{v}) = 0$. Zusätzlich liegen die Meßwerte einer Instrumentalvariablen (IV) $\underline{z}^p = \underline{z} + \underline{v}^z$ vor mit den Eigenschaften $E(\underline{v}^z, \underline{z}) = E(\underline{v}^z, \underline{v}^p) = E(\underline{v}^z, \underline{x}) = E(\underline{v}^p, \underline{z}) = E(\underline{v}^p, \underline{z}) = 0$, aber $E(\underline{z}, \underline{x}) \neq 0$. Dann liefert der bekannte Instrumentalvariablenansatz ¹⁾ $\hat{\gamma} = (\underline{z}'\underline{x})^{-1} \underline{z}'\underline{y}$ unverzerzte Schätzungen für γ . Dies sieht man für das oben eingeführte Modell sofort nach $E(\hat{\gamma}) = E\left\{ \frac{\text{cov}(\underline{z}, \underline{y}) + \text{cov}(\underline{v}^z, \underline{y}) + \text{cov}(\underline{z}, \underline{v}) + \text{cov}(\underline{v}^z, \underline{v})}{\text{cov}(\underline{z}, \underline{x}) + \text{cov}(\underline{v}^z, \underline{x}) + \text{cov}(\underline{v}^z, \underline{v}^p) + \text{cov}(\underline{v}^p, \underline{z})} \right\}$

$$= \frac{\gamma \text{cov}(\underline{z}, \underline{x})}{\text{cov}(\underline{z}, \underline{x})} = \gamma$$

Die zugehörige Schätzung für $\hat{\sigma}^2(\underline{v})$ kann aber nicht unmittelbar aus $\sum_{t=1}^T (y_t - \hat{\gamma} x_t^p)^2$ hergeleitet werden, da ja die Varianzzerlegung $\hat{\sigma}^2(\underline{y}) = \gamma^2 \hat{\sigma}^2(\underline{x}) + \hat{\sigma}^2(\underline{v})$ nicht für $\underline{x}^p$ gilt ²⁾:

$$\hat{\sigma}^2(\underline{\gamma} \underline{x}^p) = \gamma^2 \hat{\sigma}^2(\underline{x}) + \gamma^2 \hat{\sigma}^2(\underline{v}^p) \quad \text{und damit}$$

$$\hat{\sigma}^2(\underline{y} - \underline{\gamma} \underline{x}^p) = \hat{\sigma}^2(\underline{y} - \underline{x} - \underline{\gamma} \underline{v}^p) = \hat{\sigma}^2(\underline{v}) + \gamma^2 \hat{\sigma}^2(\underline{v}^p) \quad .$$

Zur Berechnung von $\hat{\sigma}^2(\underline{v})$ ist ein kleiner Umweg nötig nach $\hat{\sigma}^2(\underline{v}) = \hat{\sigma}^2(\underline{y}) - \text{cov}(\underline{x}^p, \underline{y})$. Dieser Ausdruck ist sofort für die multiple Regression zu verallgemeinern nach

$$\hat{\sigma}^2(\underline{v}) = \hat{\sigma}^2(\underline{y}) - \text{cov}(\underline{X}^p, \underline{y}) \quad , \quad \text{wobei } \underline{X}^p \text{ die Matrix aller}$$

Regressoren einschließlich der Proxies ist.

Schönfeld ³⁾ zeigt, daß der Instrumentalvariablenansatz auch dann konsistent ist, wenn in den obigen Annahmen die strengen Forderungen der wechselseitigen Unabhängigkeit im Erwartungswert durch die schwächere Forderung der Unkorreliertheit im Wahrscheinlichkeitslimes ersetzt werden.

Zu erwähnen ist, daß bei dem IV-Ansatz die "Umkehrregression" identische Ergebnisse liefert: Regressiert man $\underline{x}^p$ auf $\underline{y}$ mit

1) Vgl. z. B. Wonnacott & Wonnacott (1970) S. 154

2) Hierauf wird in vielen Lehrbüchern seltsamerweise nicht eingegangen; zum Teil findet man sogar ausgesprochen irreführende Darstellungen (vgl. z. B. Johnston (1972) die Seiten 280 und, auf diese Seite beziehend, S.283)

3) Schönfeld (1971) Band II S. 107

einem IV-Ansatz unter Verwendung von $\underline{z}^p$, so ist

$E\left(\frac{1}{\gamma}\right) = \frac{\text{cov}(z, x)}{\gamma \text{cov}(z, x)} = \frac{1}{\gamma}$. Diese Eigenschaft ist für IV-Schätzer typisch ¹⁾.

Zur Unterscheidung des IV-Ansatzes beim EIV-Modell zu dem in späteren Kapiteln behandelten IV-Ansatz bei Strukturgleichungen in interdependenten Modellen soll noch einmal das Ziel der IV-Schätzung betont werden: Zu schätzen sind die bei dem Vektor $\underline{x}$ oder bei den Vektoren aus X stehenden Koeffizienten und nicht die bei den gemessenen Vektoren $\underline{x}^p$ bzw. x^p stehenden Koeffizienten. Ziel der Schätzung ist also die Quantifizierung der Abhängigkeit der Variablen $\underline{y}$ von den Variablen X und nicht von den Variablen x^p . Dadurch ist nach dem Schätzen dieser Koeffizienten jedoch keine Schätzung für den Vektor $\underline{y}$ zu erhalten, sondern nur eine Schätzung für dessen Varianz. Die Korrelation des Residuums $\underline{y}$ mit den Regressoren ist also eine Eigenschaft der Variablen, die für die Schätzung zur Verfügung stehen, und nicht der Variablen, deren Regressionskoeffizienten man schätzen möchte. Im Gegensatz hierzu ist die Korrelation der endogenen Variablen in einer Strukturgleichung eine Eigenschaft dieser Variablen. Und für diese Endogenen sollen auch die Strukturgleichungskoeffizienten geschätzt werden, sie sind keine 'Proxies'. Dadurch sind nach dieser Koeffizientenschätzung auch Schätzungen für das "Residuum" $\underline{u}$ der Strukturgleichung möglich.

Das Hauptproblem bei der Anwendung eines IV-Ansatzes besteht vor allem darin, geeignete Instrumentalvariablen zu finden, die die geforderten Eigenschaften besitzen. Dies wird das folgende Beispiel zeigen.

1) Darauf wird noch in Kapitel 7 ausführlich eingegangen

3.5.1.1. Lagvariable als Instrumentalvariable

Der folgende Ansatz geht auf Reiersøl ¹⁾ zurück.

Gegeben sei wieder das Modell des Abschnitts 3.5.1. mit den dort angegebenen Annahmen für das Modell und für die Proxy $\underline{x}_p = \underline{p}$. Darüber hinaus werde von $\underline{x}$ angenommen, daß es einem autoregressiven Prozeß folge nach $x_t = \varrho x_{t-1} + \varepsilon_t$ mit $\varrho \neq 0$. Dann kann x_{t-1}^p als Instrumentalvariable für x_t^p dienen, wenn zusätzlich folgende Bedingungen erfüllt sind ²⁾:

$$E(\varepsilon_t' \underline{x}_{t-1}) = E(\underline{v}' \underline{x}_{t-1}) = E(\underline{v}' \underline{p}_{t-1}) = E(\underline{v}_{t-1}^p \underline{v}^p) = E(\underline{v}^p' \underline{p}_{t-1}) = 0 \quad 3)$$

wenn durch das Subskript 't' jeweils der durch Streichen des ersten Meßpunktes entstandene Vektor, durch das Subskript 't-1' jeweils der durch das Streichen des letzten Meßpunktes entstandene Vektor beschrieben wird. Dann ist

$$\text{cov}(\underline{v}_t, \underline{p}_{t-1}) = \text{cov}(\gamma \underline{x} + \underline{v}, \underline{x}_{t-1} + \underline{v}_{t-1}^p) = \gamma (\text{cov}(\underline{x}_t, \underline{x}_{t-1}) + \text{cov}(\underline{x}_t, \underline{v}_{t-1}^p))$$

$$\text{und } \text{cov}(\underline{p}_t, \underline{p}_{t-1}) = \text{cov}(\underline{x}_t, \underline{x}_{t-1}) + \text{cov}(\underline{x}_t, \underline{v}_{t-1}^p)$$

und ist abzuschätzen nach

$$E(\hat{\gamma}) = \frac{\text{cov}(\underline{p}_{t-1}, \underline{y}_t)}{\text{cov}(\underline{p}_{t-1}, \underline{p}_t)} = \gamma$$

Damit werden sofort die Grenzen dieser Instrumentalvariablenwahl deutlich, denn die Verwendung einer Lag-Variablen als IV verspricht ja nur bei großem ϱ eine effiziente Verwendung. ⁴⁾. Dies allerdings wird, vor allem bei der Verwendung von wirtschaftlichen Zeitreihendaten, wohl im allgemeinen gewährleistet sein.

Problematisch sind jedoch die Annahmen, die eine "Lagkorrelation" sowohl von $\underline{v}_t^p$ mit $\underline{v}_{t-1}^p$ wie auch von $\underline{v}_t$ mit $\underline{x}_{t-1}$ und $\underline{v}_{t-1}^p$ ausschließen. Deshalb müssen diese Annahmen diskutiert werden.

1) Reiersøl (1945 (1))

2) Vgl. z. B. Maddala (1977) S.298

3) Maddala führt hier nicht alle notwendigen Annahmen auf

4) Zur Effizienz von IV-Ansätzen vgl. z. B. Wonnacott & Wonnacott (1970) S. 159f und S. 359

a) $E(\underline{v}_t^p, \underline{v}_{t-1}^p) = 0$ oder zumindest $\text{plim}(\frac{1}{T} \underline{v}_t^p, \underline{v}_{t-1}^p) = 0$

Der Fehler in $\underline{p}$ soll also nicht autokorreliert sein. Dieser "Fehler" kann jedoch entweder ein Meßfehler oder ein Approximationsfehler sein. Nimmt man den Meßfehler als reinen Zufallsfehler an, so ist die Unkorreliertheit zeitlich benachbarter Meßfehler plausibel. Zu bezweifeln ist diese Unkorreliertheit jedoch, wenn systematische Verzerrungen, wie sie für Aggregationsfehler typisch sind, bei der Messung auftreten. Denn diese beruhen ja auf einer falschen Systematik der Messung von $\underline{x}$, und das Erkennen und Eliminieren eines solchen Fehlers zieht sich sicher über mehrere Messungen hin. Ähnlich sieht es im Fall der Verwendung einer Proxy aus ¹⁾. Da hier der Fehler ein "Abbildungsfehler" ist, ist es gerade bei Zeitreihendaten kaum zu erwarten, daß diejenigen Störeinflüsse, die die Ähnlichkeit zum Zeitpunkt t überdurchschnittlich gut oder schlecht werden lassen ²⁾, schon zum nächsten Zeitpunkt nicht oder nur noch unwesentlich nachwirken.

b) $E(\underline{v}_t^p, \underline{v}_{t-1}^p) = 0$ oder zumindest $\text{plim}(\frac{1}{T} \underline{v}_t^p, \underline{v}_{t-1}^p) = 0$

Hier ist die verschwindende "Lagkorrelation" plausibler: Da der Fehler $\underline{v}_t$ in der Regressionsbeziehung von $\underline{y}_t$ auf die $\underline{x}$ -Regressoren entweder einen Meßfehler in $\underline{y}_t$ oder einen Fehler durch das Ausschließen von Regressoren aus dieser Gleichung darstellt, der Fehler $\underline{v}_{t-1}^p$ jedoch den Meßfehler einer anderen Messung oder den "Abbildungsfehler" einer anderen Korrelationsbeziehung widerspiegelt, können diese beiden Prozesse sowohl für zeitgleiche wie für zeitverschobene Betrachtungen als unabhängige Prozesse betrachtet werden.

1) Es wird hier davon ausgegangen, daß der Ökonometriker, der diese Plausibilitätsüberlegungen anstellen muß, sich darüber im klaren ist, ob er es mit einem Meßfehler oder mit einem Approximationsfehler zu tun hat.

c) $E(\underline{v}_t^p \underline{x}_{t-1}) = 0$ oder zumindest $\text{plim}(\frac{1}{T}(\underline{v}_t^p \underline{x}_{t-1})) = 0$

Auch diese Annahme kann im allgemeinen akzeptiert werden, denn sowohl einen Meßfehler in dem fehlerbehafteten Regressor wie auch einen Abbildungsfehler zum Zeitpunkt t kann man dann als mit $\underline{x}_{t-1}$ unkorreliert ansehen, wenn weder das Meßverfahren noch die Wahl der Proxy den Schwankungen von $\underline{x}_{t-1}$ angepaßt werden. Das ist jedoch schon wegen der Unkenntnis des "wahren" Wertes von $\underline{x}_{t-1}$ unmöglich.

d) $E(\underline{v}_t \underline{x}_{t-1}) = 0$ oder zumindest $\text{plim}(\frac{1}{T}(\underline{v}_t \underline{x}_{t-1})) = 0$

Hier kann in gleicher Weise wie beim Abschnitt b) argumentiert werden. Zu beachten ist hierbei, daß der Ansatz ohne $\underline{x}_{t-1}$ als Regressor, also $\underline{y} = \underline{x}_t \underline{\gamma} + \underline{v}_t$, als der richtige Ansatz angenommen wird, also $\underline{v}_t$ kein fälschlicherweise ausgeschlossenes $\underline{x}_{t-1}$ beinhaltet.

Folgt man den obigen Argumenten, so kann man in den Annahmen $E(\underline{v}_t^p \underline{v}_{t-1}^p) = 0$ und $\sigma(\underline{x}_t, \underline{x}_{t-1}) \neq 0$ die stärksten Restriktionen für die Anwendung des beschriebenen IV-Ansatzes sehen. Aus diesem Grunde wohl bezeichnet Maddala ¹⁾ die Lagvariablen als "very poor instruments". Bejaht man allerdings die Annahme $\sigma(\underline{x}_t, \underline{x}_{t-1}) \neq 0$ und erweitert diese Modellvorstellung zu der Annahme eines autoregressiven Prozesses höherer Ordnung, so kann man, wie z. B. Novak ²⁾ zeigt, zu durchaus effizienten IV-Schätzungen gelangen. Aber gerade die Arbeit von Novak zeigt auch deutlich, mit wievielen zusätzlichen restriktiven Annahmen man operieren muß, um zu einem konsistenten Schätzer zu gelangen.

1) Maddala (1977) S. 298

2) Novak (1975a), (1975b), (1976), (1977)

3.5.1.2. Klassenzugehörigkeit als Instrumentalvariable

Um die restriktiven Annahmen zu vermeiden, die die Verwendung von Lagvariablen als Instrumentalvariablen einschränken, wurden Klassifikationsverfahren zur Schätzung von EIV-Modellen vorgeschlagen von Wald ¹⁾, Bartlett ²⁾ und Durbin ³⁾. Da alle drei Vorschläge auf der Bildung von Rangfolgen von $\underline{x}^D = \underline{p}$ beruhen, soll zur Veranschaulichung nur das Verfahren von Wald beschrieben werden.

Gegeben seien die Modellannahmen des Abschnitts 3.1.

Nach Wald werden die $x_t^D = p_t$ nach ihrer Größe aufsteigend geordnet und alle Wertepaare y_t, p_t der Gruppe I zugeordnet, für die der Wert von p_t kleiner als der Median m der Werte $\underline{p}$ ist. Die übrigen Wertepaare werden der Gruppe II zugeordnet. Als Schätzung für γ dient dann der Ausdruck

$$\hat{\gamma}_{\text{Wald}} = (\bar{y}_{\text{II}} - \bar{y}_{\text{I}}) / (\bar{p}_{\text{II}} - \bar{p}_{\text{I}}) = \Delta \bar{y} / \Delta \bar{p}$$

wobei die $\bar{y}$ bzw. $\bar{p}$ die jeweiligen arithmetischen Mittel der Klassen I und II für $\underline{y}$ und $\underline{p}$ sind. Diesen Ansatz kann man als einen IV-Ansatz betrachten, wobei die Instrumentalvariable für alle zur Gruppe I gehörenden Messungen den Wert -1, für die übrigen Werte den Wert +1 erhält. Dem Ansatz liegt die Vorstellung zugrunde, daß

$$\bar{p}_{\text{I}} = \frac{1}{T_{\text{I}}} \sum_{t=1}^{T_{\text{I}}} p_{\text{I},t} = \frac{1}{T_{\text{I}}} \sum x_{\text{I},t} + v_{\text{I},t} \approx \bar{x}_{\text{I}} \quad (1)$$

$$\bar{y}_{\text{I}} = \frac{1}{T_{\text{I}}} \sum_{t=1}^{T_{\text{I}}} y_{\text{I},t} = \frac{1}{T_{\text{I}}} \sum \gamma x_{\text{I},t} + v_{\text{I},t} = \gamma \bar{x}_{\text{I}} \quad (2)$$

und entsprechend $\bar{p}_{\text{II}} \approx \bar{x}_{\text{II}}$ bzw. $\bar{y}_{\text{II}} = \gamma \bar{x}_{\text{II}}$ gilt.

Während aber (2) wegen $E(v/x) = 0$ im Erwartungswert erfüllt ist, gilt dies für (1) nicht. Denn durch die Eingruppierung

1) Wald (1940 (1))

2) Bartlett (1949)

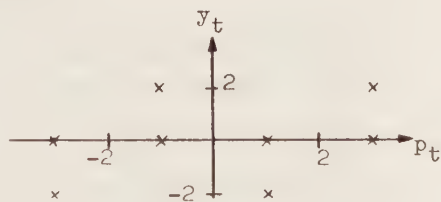
3) Durbin (1954)

der Wertepaare nach der Größe von p_t ist diese Klassifikation mit $\underline{v}^p$ korreliert: Alle Werte $p_t = x_t + v_t^p$, für die $v_t^p < m - x_t$ ist, werden fälschlicherweise der Gruppe I zugeordnet; ebenso werden alle Werte p_t , für die $v_t^p \geq m - x_t$ ist, fälschlicherweise der Gruppe II zugeordnet. Bezeichnet man nun die falsch zugeordneten Werte mit p_I^f bzw. p_{II}^f , die zugehörigen Fehler mit $v_I^{p,f}$ und $v_{II}^{p,f}$, so sieht man sofort, daß jetzt $E(\Delta v_I^p) = 0 - E(v_I^{p,f}) - E(v_{II}^{p,f}) = -E(\Delta v_{II}^p)$ ist.

Ein Extremfall soll dies verdeutlichen. Für das Modell

$\underline{y} = 1 \cdot \underline{x} + \underline{v}$ wird mit der Meßreihe

x_t	-1	-1	-1	-1	+1	+1	+1	+1
v_t^p	-2	-2	+2	+2	-2	-2	+2	+2
v_t	-1	+1	-1	+1	-1	+1	-1	+1
p_t	-3	-3	+1	+1	-1	-1	+3	+3
y_t	-2	0	-2	0	0	+2	0	+2



der Schätzer $\hat{\gamma}_{\text{Wald}} = (0 - 0) / (2 + 2) = 0$ erhalten, da die Differenz $\bar{p}_{II} - \bar{p}_I$ jetzt nur noch die Summe der $v^{p,f}$ widerspiegelt, die Klassifizierung also nicht nach x_t , sondern nach p_t durchgeführt wurde.

Damit verletzt der Wald'sche Schätzer die für Instrumentalvariablen notwendige Bedingung $E(\underline{v}^p | \underline{z}) = 0$, da ja die Klasseneinteilung mit $\underline{v}^p$ korreliert ist. Nach Neyman und Scott ¹⁾ kann man mit dem Wald'schen Schätzer eine konsistente Schätzung nur erhoffen, wenn $\underline{p}$ so in zwei Gruppen aufzuspalten ist, daß die Wahrscheinlichkeit $w(v_I^{p,f} \neq 0) = 0$ ist; dies bedeutet anschaulich aber nichts anderes als die Forderung nach einer ausreichend großen "Lücke" zwischen den beiden Gruppen $\underline{x}_I$ und $\underline{x}_{II}$, daß auch durch eine Überlagerung mit den Fehlern $\underline{v}_I^p$ und $\underline{v}_{II}^p$ keine Überlappung der zugehörigen Gruppen $\underline{p}_I$ und $\underline{p}_{II}$ erfolgen kann: Die beiden 'Cluster' müssen deutlich getrennt sein.

1) Neyman & Scott (1952)

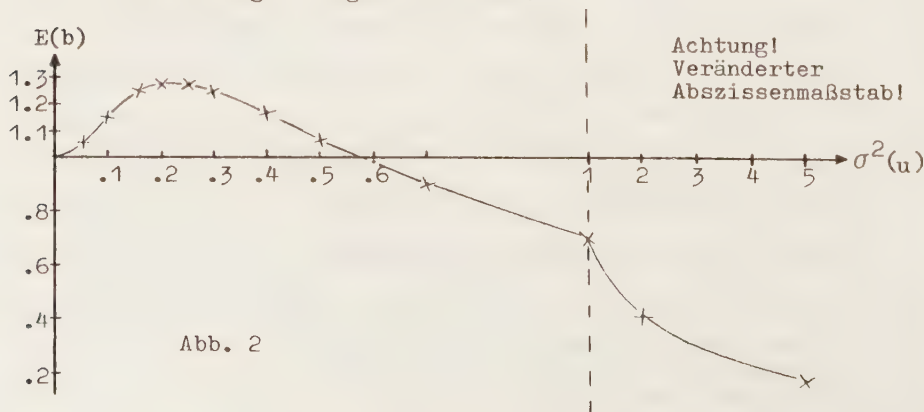
Der nun naheliegende Gedanke, ohne eine Gruppierung nach der Größe der p_t zwei Gruppen von jeweils dem Umfang $T_I = T_{II}$ mit $T_I + T_{II} = T$ durch zufälliges Zuordnen zu bilden, führt aber auch nicht zum Ziel. Denn wegen $E(\underline{v}^P, \underline{v}) = 0$ gilt mit $\Delta \bar{x} = \bar{x}_{II} - \bar{x}_I$ und $\Delta \bar{v}^P = \bar{v}_{II}^P - \bar{v}_I^P$ zwar dann $E(\hat{\gamma}^*) = E((\bar{y}_{II} - \bar{y}_I)/(\bar{p}_{II} - \bar{p}_I)) = E(\gamma \Delta \bar{x} / (\Delta \bar{x} + \Delta \bar{v}^P)) = \gamma \Delta \bar{x} \cdot E(\frac{1}{\Delta \bar{x} + \Delta \bar{v}^P})$, dies ist aber nicht $\gamma \Delta \bar{x} / \Delta \bar{x} = \gamma$.

Zur Quantifizierung der Verzerrung dieses Ansatzes wurden mit numerischer Integration an einer mit 11 signifikanten Ziffern rechnenden EDV-Anlage die Erwartungswerte des Ausdrucks $b = \frac{1}{1+u}$ durchgerechnet, wobei u als $(0; \sigma^2)$ -normalverteilt angenommen wurde und die Varianz von u von Null aufsteigend variiert wurde; damit hat man $E(\hat{\gamma}^*) = \gamma \cdot b$ in Abhängigkeit des Verhältnisses der Varianz $\sigma^2(u)$ zu $(\Delta \bar{x})^2$. Die Unendlichkeitsstelle für $u = -1$ wurde durch Ausklammern des Bereichs $-1.0 - 1 \cdot 10^{-9} \leq u \leq -1.0 + 1 \cdot 10^{-9}$ umgangen. Dies ergab folgende Ergebnisse

$\sigma^2(u)$	0.01	0.02	0.04	0.05	0.1	0.16	0.2	0.25	0.3
E(b)	1.010	1.021	1.046	1.062	1.157	1.254	1.281	1.279	1.253

$\sigma^2(u)$	0.4	0.5	0.585	0.6	0.7	1.0	2.0	5.0
E(b)	1.169	1.075	1.000	0.987	0.908	0.724	0.424	0.187

die in Abbildung 2 dargestellt sind:



Diese Ergebnisse erschienen dem Verfasser aus mehreren Gründen so bemerkenswert, daß er sie nicht unterschlagen mochte, auch wenn die zugrundeliegende Rechnung trivial ist. Denn aus dem Verlauf von $E(b) = f(\sigma^2(u))$ kann man erkennen:

1. Die Richtung der Verzerrung eines solchen "Quotientenschätzers" ist nicht unabhängig von $\sigma^2(u)$.
2. Besteht die 'Varianz' des Nenners nur aus der Varianz von u , d. h. $\sigma^2(u) \gg (E(\Delta x))^2$, dann ist der Quotientenschätzer gegen Null verzerrt: Ein solcher Fall liegt ja bei dem Extrembeispiel der S. 35 vor.
3. Da die Varianz $\sigma^2(u)$ der obigen Rechnung für $\sigma^2(\bar{v}^p)$ steht, also die Varianz der Differenz der "Klassenfehlermittel", geht für $T \rightarrow \infty$ auch $\sigma^2(\bar{v}^p) = \sigma^2(u) = \frac{4}{\pi} \sigma^2(v^p)$ gegen Null. Hat man also eine von p bzw. von $\bar{v}^p$ unabhängige Klassifikation, so ist der Schätzer $\gamma^{\oplus}$ zwar verzerrt, aber konsistent.

Eine nach Punkt 3. von $\bar{v}^p$ unabhängige Klassifikationsregel hat man aber z. B. bei Zeitreihendaten, wenn diese einem Trend folgen und die Varianz von $\bar{v}^p$ homoskedastisch ist: Ordnet man nämlich die Wertepaare $(p_t; y_t)$ für $t \leq \frac{1}{2} T$ in die erste Klasse, die restlichen Wertepaare in die zweite Klasse, so ist für $E(\Delta \bar{p}) = E(\Delta \bar{x} + \Delta \bar{v}^p)$ ein Wert von $\Delta \bar{x} \neq 0$ zu erwarten, wenn der Trend (gleichgültig ob linear oder nicht) zunimmt. Für einen positiven Wert von $\Delta \bar{x}$ aber hat man nach 3. einen konsistenten Schätzer.¹⁾

Mindestens ebenso interessant ist aber die Quantifizierung der Verzerrung in Abhängigkeit von $\sigma^2(u)$, da der Quotientenschätzer eine von der OLS-Verzerrung abweichende Abhängigkeit zeigt.

1) Die Annahme eines wachsenden, aber beliebigen Trends in $\bar{x}$ ist nicht zu verwechseln mit den Annahmen eines autoregressiven Prozesses in $\bar{x}$ des Abschnitts 3.5.1.1. ; auch muß über die Form des Trends nichts bekannt sein.

Denn bei einer nicht zu kurzen Zeitreihe vom Umfang T besitzt man nun ein, wenn auch grobes, Kontrollinstrument bezüglich der Verzerrung des Quotientenschätzers, wenn man nach dem folgenden Verfahren vorgeht:

Man teile den gesamten Stichprobenumfang T in z. B. 8 gleichgroße Klassen ein, wobei der laufende Index das Klassifikationsmerkmal sei; somit stellen z. B. die Messungen für $t=1, \dots, 3$ die erste Klasse, diejenigen für $t=4, \dots, 6$ die zweite Klasse und schließlich bilden die Messungen mit $t=22, \dots, 24$ die letzte Klasse. Hiermit kann man $\frac{8 \cdot 7}{2} = 28$ Quotientenschätzungen $\hat{y}_i^{\oplus}$ ($i=1, \dots, 28$) berechnen und aus ihnen das arithmetische Mittel bilden, das eine Schätzung für $E(\hat{y}^{\oplus} / T=3)$ darstellt. Anschließend teilt man die Messungen nach der beschriebenen Zuordnungsregel in 6 gleichgroße Klassen ein und erhält mit dem arithmetischen Mittel der insgesamt $\frac{6 \cdot (6-1)}{2} = 15$ Quotientenschätzer $\hat{y}_i^{\oplus}$ ($i=1, \dots, 15$) eine weitere Schätzung für $E(\hat{y}^{\oplus} / T=4)$. Eine dritte Aufteilung des Stichprobenumfangs, jetzt aber in 4 Klassen, ergibt mit dem arithmetischen Mittel der 6 möglichen Quotientenschätzer eine weitere Schätzung für einen bedingten Erwartungswert von $E(\hat{y}^{\oplus} / T=6)$. Diese drei bedingten Erwartungswerte liegen bei Fehlen von Autokorrelation und Heteroskedastizität im Erwartungswert auf drei Punkten, die das γ -fache dreier Punkte der Abbildung 2 darstellen, wobei mit steigenden T nach $\sigma^2(u) = \sigma^2(\bar{v}^D) = \frac{4}{T} \sigma^2(v^D)$ die Varianz von u kleiner wird. Hätte man zusätzlich die Werte von $(\Delta \bar{x})^2$, so könnte man durch einen Vergleich der Steigungen der zwischen den drei Punkten liegenden Änderungen von $E(\hat{y}^{\oplus})$ in Abhängigkeit von T das Ausmaß der Verzerrung bestimmen. Ohne diese Kenntnis hat man bei einem steigenden, aber sonst beliebigen Trend lediglich die Möglichkeit, bei einem Umkippen der Steigungen,

wie es bei Abbildung 2 in der Umgebung des Bereiches für $\sigma^2(u) \approx 0.2$ der Fall ist, einen Bereich für γ anzugeben nach $E(\hat{\gamma}^{\oplus}/T = 6) / 1.28 \leq \gamma \leq E(\hat{\gamma}^{\oplus}/T=6)$. Wird bei wachsendem T $E(\hat{\gamma}^{\oplus})$ kleiner, kann man den mit dem größten T geschätzten Wert $E(\hat{\gamma}^{\oplus})$ als den besten Schätzer auswählen. Lediglich bei wachsendem T und größer werdendem $E(\hat{\gamma}^{\oplus})$ ist keinerlei Eingrenzung der Verzerrung möglich, da man ja nicht weiß, in welchem Bereich 'rechts' der waagerechten Tangente der Figur 2 man sich befindet.

3.6. Bewertung der Lösungsansätze

Die Schwächen der gezeigten Lösungsansätze sind offenbar; man sieht deutlich, daß entweder der Lösungsansatz eine Fülle von Annahmen benötigt ¹⁾ oder bei Beschränkung der nötigen Annahmen nur verhältnismäßig eingeschränkte Anwendung finden kann. ²⁾ Alle Lösungsansätze haben aber gemeinsam, daß sie die zusätzlichen Informationen, die für eine zumindest konsistente Schätzung der Regressionskoeffizienten nötig sind, durch die Annahme von Beziehungen der in der zu schätzenden Gleichung vorhandenen Variablen gewinnen: Es sind Schätzer für Eingleichungssysteme, die deshalb bei Lösungsansätzen nur auf die "Binnenstrukturen" dieser Gleichungen bzw. der in ihnen enthaltenen Variablen zurückgreifen können. In Kapitel 6 werden wir sehen, daß beim Vorliegen von Mehrgleichungssystemen durch die Verwendung von Annahmen über die Beziehungen zwischen den Gleichungen unter Umständen effektivere zusätzliche Informationen zur Lösung des Schätzproblems eingesetzt werden können.

1) So benötigt z. B. Novak (1975) insgesamt 14 zum Teil sehr restriktive Annahmen!

2) Alle Schätzer erlauben die Existenz lediglich einer fehlerbehafteten Variablen; Mehrregressormodelle mit einer fehlerbehafteten können durch entsprechende Bereinigungsverfahren (Schneeweiß (1974) S. 141) auf die bei den Lösungsansätzen behandelten Einregressormodelle zurückgeführt werden.

3.7. Kritische Modelle

Wie in dem folgenden Kapitel noch verdeutlicht wird, ist die entscheidende Ursache für eine inkonsistente OLS-Schätzung einer Regressionsgleichung die Korreliertheit von Regressoren und Fehlerterm der Regression. Deshalb ist schon bei der Formulierung eines zu schätzenden Regressionsansatzes darauf zu achten, daß weder das Ausgangsmodell noch eventuelle Transformationen dieses Modells mit dem Fehlerterm korrelierte Regressoren besitzen oder erzeugen. Daß diese auf S. 4 veranschaulichte Annahme 2 in der ökonometrischen Schätzpraxis leicht unbemerkt verletzt wird, sollen ohne Anspruch auf Vollständigkeit folgende Beispiele zeigen.

3.7.1. Modelle mit verzögerten endogenen Variablen als Regressoren und autokorrelierten Fehlern

Für den Ansatz

$$y_t = \sum_{i=1}^{k1} x_{it} \gamma_i + \sum_{i=k1+1}^{k2} y_{i,t-\tau} \gamma_i + v_t \quad (1)$$

mit

$$v_t = \sum_{\tau=1}^s v_{t-\tau} \delta_{\tau} + \varepsilon_t \quad (2)$$

ist zumindest dann, wenn sich unter den $y_{i,t-\tau}$ auch der Regressand $y_{t-\tau}$ befindet, wegen der Korreliertheit von $y_{t-\tau}$ und $v_{t-\tau}$ nach der Beziehung

$$r_{y,v} = \frac{\text{cov}(y;v)}{\sqrt{\text{var}(y) \cdot \text{var}(v)}} = \frac{\sqrt{\text{var}(v)}}{\sqrt{\text{var}(y)}} = \sqrt{1 - R^2} \quad (3)$$

auch $y_{t-\tau}$ und v_t wegen der Beziehung (2) korreliert.

Diese Korrelation wird zwar wegen des gerade bei Zeitreihenmodellen üblichen großen R^2 keine großen Werte annehmen, außerdem kann ja wegen des "proximity theorems" von Wold ¹⁾ die Verzerrung der geschätzten Regressionskoeffizienten

1) Vgl. S. 6 Abschnitt 1.1.

dann vernachlässigt werden. Vorsicht ist aber unbedingt geboten, wenn bei der Schätzung einer Zeitreihenregression erst durch die Hereinnahme verzögerter Endogener das Bestimmtheitsmaß der Regression hohe Werte annimmt und der Varianzanteil der Exogenen an der durch die Regressions-schätzung erklärten Varianz gering wird.

3.7.2. Durch Transformation in absolute Differenzen entstandene Modelle

Um die durch Kollinearitäten zwischen den Regressoren auftretende Vergrößerung der Varianzen der geschätzten Regressionskoeffizienten zu beseitigen oder zumindest zu mildern, werden in der Schätzpraxis bisweilen statt der gemessenen Werte der Variablen deren Differenzen in das Schätzmodell eingebracht ¹⁾. Geschätzt wird also statt des Modells

$$y_t = \sum_{i=1}^K x_{it} \gamma_i + v_t \quad \text{das transformierte Modell}$$

$$\Delta y_t = \sum_{i=1}^K \Delta x_{it} \gamma_i + \Delta v_t = y_t - y_{t-1} = \sum_{i=1}^K (x_t - x_{t-1}) \gamma_i + v_t - v_{t-1}$$

Schneeweiß ²⁾ weist schon auf einen möglichen Effizienzverlust durch eine solche Transformation hin, da bei Vorliegen eines Ausgangsmodells mit nicht autokorrelierten v_t die Transformation autokorrelierte Residuen erzeugt; nach

$$\text{cov}(\Delta v_t; \Delta v_{t-1}) = \text{cov}((v_t - v_{t-1}); (v_{t-1} - v_{t-2})) =$$

$$E(v_t v_{t-1} - v_{t-1}^2 - v_t v_{t-2} + v_{t-1} v_{t-2}) = -\sigma^2(v) \quad (1)$$

und

$$\text{var}(\Delta v_t) = E(v_t - v_{t-1})^2 = 2 \cdot \sigma^2(v_t) = 2\sigma^2(v) \quad (2)$$

ergibt sich ein Autokorrelationskoeffizient $\rho = -0.5$ ³⁾.

Damit ist aber OLS zwar noch ein unverzerrter Schätzer,

1) Vgl. z. B. Schneeweiß (1974) S. 143 f.

2) Schneeweiß (1974) S. 143

3) Man beachte, daß damit die Schätzung $s^2 = \frac{1}{T-K-1} \Delta y' \Delta y$ keine unverzerrte Schätzung für $\text{var}(\Delta v)$ liefert; vgl. S. 7, Abschnitt 1.1.

aber eben nicht mehr "BLUE".

Eine mögliche Quelle für eine Inkonsistenz des OLS - Schätzers ist aber auch hier durch eine mögliche Korrelation der Regressoren mit dem Residuum der Regressionsgleichung gegeben. Denn selbst wenn in dem Ausgangsmodell die Residuen mit den Regressoren unkorreliert waren, bleibt für das transformierte Modell noch der Term

$$\text{cov}(\Delta x_{it}; \Delta v_t) = - E(x_{i,t-1} v_t + x_{it} v_{t-1}) \quad (3)$$

übrig. Dieser Ausdruck ist aber ohne zusätzliche Annahmen nicht zwangsläufig Null; er ist es vor allem dann nicht, wenn im Modell verzögerte Endogene vorhanden sind, also $x_{jt} = y_{t-1}$. Denn dann wäre $\text{cov}(\Delta x_{jt} \Delta v_t)$ nur dann Null, wenn

$$E(x_{j,t-1} v_t) = - E(x_{jt} v_{t-1}) \quad (4)$$

und damit

$$E(y_{t-2} v_t) = - E(y_{t-1} v_{t-1}) \neq 0 \quad (5)$$

ist für $\text{var}(v) \neq 0$. Da aber nach Schneeweiß ein Differenzenansatz besonders bei positiv autokorrelierten v_t einen Effizienzgewinn bedeutet, ist dann eine Erfüllung der Bedingung (5) ausgeschlossen, da bei einem Autokorrelationskoeffizienten des ursprünglichen Modells $\rho = +1$ jetzt

$$\text{cov}(\Delta v_t \Delta v_{t-1}) = E(v_t v_{t-1} - v_{t-1}^2 - v_t v_{t-2} + v_{t-1} v_{t-2}) = 0$$

ist.

3.7.3. Regelungsmodelle (Rückkopplungsmodelle)

Ein wichtiges Einsatzgebiet ökonomischer Modelle ist die Auswahl der "richtigen" Steuerungspolitik ¹⁾, d. h. wie groß müssen die Werte der Exogenen gewählt werden, um bestimmte Zielgrößen der Endogenen zu erreichen. Daß in der Praxis aber nicht nur eine Steuerung, sondern eine "Regelung" ²⁾

1) Vgl. z. B. Koutsoyiannis (1973) S. 8

2) "Regelung" ist im kybernetischen Sprachgebrauch Steuerung mit Rückkopplung

vorliegen kann, soll an einem kleinen Beispiel gezeigt werden.

Nehmen wir an, daß für eine regelnde Instanz (z. B. eine Zentralbank) ein einziges Instrument zur Verfügung steht, das in einem Modellansatz auch als Regressor dient; es sei

$$y_t = \sum_{i=1}^{K-1} x_{it} \gamma_i + x_{Kt} \gamma_K + v_t \quad (1)$$

wobei die x_i ($i=1, \dots, K-1$) die nicht beeinflussbaren Exogenen und x_K das zur Regelung benutzte Instrument darstellen soll. Nehmen wir weiter an, $\eta_t = E(y_t)$ sei die angestrebte Zielgröße des Eingriffs der regelnden Instanz. Dann kann man zwei grundsätzliche Regelungsmechanismen unterscheiden:

a) Regelung ohne Zeitverzögerung (Rückkopplung ohne Totzeit)

Um y_t auf den Wert der Zielgröße η_t zu bringen, muß auch der Fehler v_t "ausgeregelt" werden. Das bedeutet, daß mit x_{Kt}^R als der für die Ausregelung des Fehlers benutzten Steuergröße

$$v_t - x_{Kt}^R \gamma_K = 0 \quad (2)$$

sein muß. Als Folge einer idealen Ausregelung des Fehlers v_t würde also die zur Beeinflussung benutzte Variable x_K zu $x_{Kt}^{\bar{R}} - x_{Kt}^R \gamma_K$ gewählt werden, d. h. jetzt sind v_t und x_{Kt} korreliert, da mit $x_{Kt}^{\bar{R}}$ als der vor der Regelung festgelegten Höhe von x_{Kt} sich nun x_{Kt} zusammensetzt nach

$$x_{Kt} = x_{Kt}^{\bar{R}} - x_{Kt}^R \gamma_K \quad (3)$$

und ja x_{Kt}^R wegen (2) mit v_t korreliert ist; x_K und y haben ihre Rollen vertauscht, denn jetzt ist ja x_K Träger des Regressionsfehlers v , während der Fehler von y ausgeregelt wurde ¹⁾; es liegt der Fall "Fehler in dem Regressor" vor.

1) Ein numerisches Beispiel ist im Anhang Seite A. III.

Die auf den ersten Blick nicht plausible Annahme der vollständigen und unverzögerten Ausregelung des Fehlers ist jedoch auf dem Gebiet gerade der Wirtschaftsdaten nicht selten, da hier neben Zustandsdaten auch kumulierte Flußgrößen verwendet werden wie z. B. über einen Zeitraum aufgelaufene Umsätze, Geld-, Kapital- bzw. Warentransfers oder getätigte Investitionen. Dann ist jedoch z. B. eine Steuerungsinstanz, die die täglichen Transferumfänge kennt, sicher zu einer Ausregelung imstande, die die Zielgröße "gesamter Transfer im Monat" sehr genau erreicht. Für den außerhalb dieser Instanz stehenden Ökonometriker ist bei seiner Schätzung der Regressionskoeffizienten unter Verwendung von Monatsdaten die Regelung aber unendlich schnell und vollkommen, wobei die Vollkommenheit ja keine Bedingung für eine Korrelation von x_{Kt} und v_t ist.

b) Regelung mit Zeitverzögerung

Bei der Ausregelung von Fehlern v_t wird in vielen Fällen, auf jeden Fall aber bei der Regelung von Zustandsgrößen, der Regelvorgang entweder erst zum nächsten Zeitpunkt¹⁾ wirken ("Regelung mit Totzeit") oder erst zum nächsten Zeitpunkt eingesetzt werden können. Dann aber gehen in die Regelgröße auch die Erwartungen ein, welche Werte die nicht beeinflussbaren Größen $x_{i,t+1}$ ($i=1, \dots, K-1$) und der auszuregelnde Fehler v_{t+1} annehmen werden. Damit liegt ein Modell vor, das in der angelsächsischen Literatur mit PAAE - Modell = partial adjustment/Adaptive expectations - Modell bezeichnet wird.²⁾ Solch eine Regelung hat aber nur dann einen Sinn, wenn über den Fehler v_{t+1} Aussagen gemacht werden können; dies

1) Dieser Zeitpunkt kann, muß aber nicht mit dem nächsten zur Schätzung verwendeten Meßpunkt übereinstimmen.

2) z. B. Johnston (1972) S. 300; Dhrymes (1971)

wird im allgemeinen über die Annahme eines autoregressiven Prozesses in v vorgenommen werden, daß also

$$v_t = \sum_{\tau=1}^s v_{t-\tau} \delta_{\tau} + \varepsilon_t \quad \text{ist. Somit zieht jede}$$

Regelung, die ihre Regelgröße nach v_t ausrichtet, wegen $x_{Kt}^R = f(v_{t-1})$ und der Existenz des autoregressiven Prozesses in v eine Inkonsistenz des OLS - Schätzers nach sich.

Einen Beweis hierfür findet man bei Doran und Griffiths ¹⁾.

Die in den letzten drei Abschnitten vorgeführten "kritischen Modelle" sollten keineswegs einen enumerativen Katalog aller Modelle ergeben, bei denen unter Anwendung von OLS inkonsistente Schätzungen zu erwarten sind. Sie stellen lediglich warnende Beispiele dar, wie leicht eine Inkonsistenz des OLS-Schätzers entstehen kann, selbst wenn das zugrundegelegte Modell keine auf den ersten Blick erkennbare Korrelation der Regressoren mit dem Residuum der Regressionsschätzung aufweist.

1) Doran & Griffiths (1978)

§ 4. Gemeinsamkeiten von 'unterspezifiziertem' und
=====
'Fehler-in-dem-Regressor' - Regressionsansatz
=====

Wie die beiden vorhergehenden Kapitel zeigten, führt sowohl ein 'unterspezifizierter' wie auch ein 'EIV'-Regressionsansatz bei der Verwendung von OLS zu inkonsistenten Schätzungen der Regressionsparameter. Beide Fälle kann man als Folgen mangelnden Wissens des schätzenden Ökonometrikers beschreiben. Während dieser bei einem unterspezifizierten Modell zwar weiß ¹⁾, daß in dem Schätzansatz mindestens ein Regressor fehlt, diesen Mangel aber durch keinerlei Zusatzinformationen beheben kann, besitzt er im Falle eines EIV-Modells wenigstens eine Approximation für diesen fehlenden Regressor. Entsprechend müssen die zusätzlichen Informationen, die für einen zumindest konsistenten Schätzansatz nötig sind, von unterschiedlicher Qualität oder Quantität sein. Dies wird das Kapitel 6 noch verdeutlichen.

In der Praxis hat sich aber neben gezielten Lösungsansätzen eine Vorgehensweise eingebürgert, die ebenfalls zu einer Verbesserung des Schätzansatzes führen kann: Man führt versuchsweise in das Schätzmodell einen Satz von möglicherweise die abhängige Variable beeinflussenden 'Versuchsvariablen' ein und wählt nach einem oder mehreren unterschiedlich spezifizierten Schätzdurchläufen nach bestimmten 'Optimalitätskriterien' einen dieser Schätzansätze als den 'besten' aus. Solche 'Optimalitätskriterien' wie auch eine Rechtfertigung dieses Vorgehens liefern die folgenden zwei Theoreme ²⁾.

1) Von diesem Wissen muß ausgegangen werden, da ja sonst kein Anlaß zu 'Verbesserungsversuchen' besteht.

2) Bei der Verwendung der beiden Theoreme werden diese allerdings in den seltensten Fällen explizit erwähnt.

Theorem 1 ¹⁾: Bei einer gegebenen abhängigen Variablen $\underline{y}$ und zwei gegebenen Matrizen unabhängiger Variablen X und X_0 , die die Dimensionen $T \cdot K$ bzw. $T \cdot K_0$ besitzen und für die $X \neq X_0$ gilt, sei das wahre Modell $\underline{y} = X\underline{\gamma} + \underline{v}$ mit $E(\underline{v}) = \emptyset$ und $\sigma^2(\underline{v}) = \sigma^2 \cdot I$. Führt man mit den beiden alternativen Matrizen X bzw. X_0 jeweils eine OLS-Schätzung durch und bezeichnet man die damit erhaltenen Schätzungen für $\sigma^2(\underline{v})$ mit $s^2 = \frac{1}{T-K} \underline{y}' M \underline{y}$ bzw. $s_0^2 = \frac{1}{T-K_0} \underline{y}' M_0 \underline{y}$ mit $M = I - X(X'X)^{-1}X'$ bzw. $M_0 = I - X_0(X_0'X_0)^{-1}X_0'$, dann gilt bei nichtstochastischen Regressoren in X und X_0

$$E(s_0^2 - s^2) \geq 0$$

Denn dann gilt $E(s^2) = \sigma^2$ und ²⁾

$$E(\underline{y}' M_0 \underline{y}) = \underline{\gamma}' X' M_0 X \underline{\gamma} + (T - K_0) \sigma^2 = (T - K_0) \sigma^2$$

Das heißt in Worten, daß der Erwartungswert für s^2 eines richtig spezifizierten Modells nicht größer werden kann als der Erwartungswert von s_0^2 eines falsch spezifizierten Modells.

Theorem 2 ³⁾: Durch das Ausschließen einer Variablen aus einer OLS-Schätzgleichung wächst das adjustierte Bestimmtheitsmaß ⁴⁾ $\bar{B}$ dann und nur dann, wenn das Quadrat des t-Wertes (also der F-Wert) des Regressionskoeffizienten dieser Variablen kleiner als 1 ist.

Da üblicherweise das adjustierte Bestimmtheitsmaß als Maß für die Güte einer Regressionsschätzung verwendet wird, besagt das Theorem 2, daß die Güte eines Regressionsansatzes durch Ausschluß eines Regressors nur dann verbessert werden kann,

1) Dieses Theorem wurde zuerst von Theil (1971) S. 543 formuliert und ist deshalb auch als "Theil's Theorem" bekannt.

2) Zum Nachvollziehen des Beweises sind folgende Aussagen hilfreich: (Vgl. z. B. Schneeweiß (1974) S. 339 u. S. 344): $E(sp(X)) = sp(E(X))$; $sp(BC) = sp(CB)$

3) Leamer (1978) S. 75

4) Das adjustierte Bestimmtheitsmaß ist definiert durch

$$1 - \bar{B} = \frac{s^2}{s_y^2} = \frac{\sum \hat{v}_t^2 / (T - K)}{\sum (y_t - \bar{y})^2 / (T - 1)}$$

wenn der F-Wert des Regressionskoeffizienten dieses Regressors kleiner als 1 ist.

Diese beiden Theoreme sind die Grundlage der bekannten schrittweisen Spezifikationsverfahren, die in praktisch jedem Statistik-Softwarepaket enthalten sind und sich auch der häufigen Verwendung gerade statistisch nicht vorgebildeter Benutzer erfreuen. Diese "extensive" Benutzung stellt auch Theil ¹⁾ fest, er warnt jedoch auch deutlich vor einem Mißbrauch. Dazu stellt er fest:

- a) Das Theorem 1 trifft nur dann zu, wenn eine der beiden alternativen Spezifikationen die richtige ist; es gilt also nicht, wenn beide Spezifikationen falsch (z. B. "nearly correct" im Sinne Liu's) sind.
- b) Das Theorem 1 liefert lediglich Aussagen über die Erwartungswerte von s^2 und s_0^2 ; deshalb ist zu erwarten, daß bei einem 'Durchspielen' vieler alternativer Spezifikationen und Auswahl derjenigen mit dem kleinsten s^2 diese Schätzung s^2 im Durchschnitt deutlich unter σ^2 liegen wird.
- c) Das Theorem 1 ist nur anwendbar, wenn die abhängige Variable für alle alternativen Spezifikationen dieselbe ist. Unzulässig wären also z. B. der Vergleich eines in $\underline{y}$ linearen Ansatzes mit einem Ansatz, bei dem $\underline{y}^+$ durch Logarithmieren oder durch Differenztransformation ("Saisonbereinigung") erhalten wurde. Das gleiche gilt, wenn $\underline{y}$ eine Linearkombination mehrerer abhängiger Variablen ist; dann muß für eine Anwendung des Theorems 1 diese Linearkombination konstant bleiben ²⁾.

Deshalb schließt Theil seine Diskussion des Theorems 1 mit der Feststellung ab, daß ein "wahrer Test (einer Spezifikation; Anm. d. Verf.) aus einer Prognose mit einem Datensatz

1) Theil (1971) S. 544 f.

2) Dieser Fall wird noch in Kapitel 7 ausführlich untersucht

besteht, der unabhängig ist von dem zur Schätzung verwendeten".

Speziell für den Fall eines fehlerbehafteten Regressors liefert das folgende Theorem ¹⁾ einen statistischen Hintergrund, das nach Wissen des Verfassers noch nicht veröffentlicht wurde.

'Theorem' 3: Stehen in einer OLS-Schätzung des wahren Modells

$\underline{y} = \underline{X} \underline{\beta} + \underline{v}$ für einen der Regressoren $\underline{x}_1$ aus $\underline{X}$ neben dem Vektor $\underline{x}_1$ eine Proxy-Variable $\underline{p}_1 = \underline{x}_1 + \underline{v}^p$ zur Verfügung, wobei $\underline{v}^p$ sowohl mit $\underline{v}$ wie auch mit den $\underline{x}_j$ ($j=1, \dots, K$) im Erwartungswert unkorreliert sein soll, dann gilt für eine OLS-Schätzung unter gleichzeitiger Verwendung von $\underline{x}_1$ und $\underline{p}_1$ nach dem Schätzansatz $\hat{\underline{y}} = \underline{X} \hat{\underline{\beta}} + \underline{p}_1 \hat{\beta}^+$

$$E(\hat{\beta}^+) = 0 \quad :$$

Die exakte Variable $\underline{x}_1$ 'verdrängt' also die Proxy-Variable vollständig.

Der Beweis sei zunächst an einem Modell mit nur zwei wahren Regressoren durchgeführt, bei dem aber zur Verdeutlichung des Schätzverhaltens für den zweiten Regressor zwei Proxies in den Schätzansatz einbezogen werden. Dann gilt:

Für das Modell $\underline{y} = \underline{x}_1 \beta_1 + \underline{x}_2 \beta_2 + \underline{v}$ mit den beiden Proxies $\underline{p}_2^{(1)} = \underline{x}_2 + \underline{v}^{p(1)}$ und $\underline{p}_2^{(2)} = \underline{x}_2 + \underline{v}^{p(2)}$ und $E(\underline{v}' \underline{v}^{p(1)}) = E(\underline{v}^{p(1)}' \underline{v}^{p(2)}) = E(\underline{x}_j' \underline{v}^{p(1)}) = E(\underline{x}_j' \underline{v}^{p(2)}) = 0$

liefert die OLS-Schätzung nach

$$(\hat{\beta}_1, \hat{\beta}_2^{(1)}, \hat{\beta}_2^{(2)})' = ((\underline{x}_1, \underline{p}_2^{(1)}, \underline{p}_2^{(2)})' (\underline{x}_1, \underline{p}_2^{(1)}, \underline{p}_2^{(2)}))^{-1} (\underline{x}_1, \underline{p}_2^{(1)}, \underline{p}_2^{(2)})' \underline{y}$$

nach etwas Rechenarbeit und mit den Bezeichnungen

$$\text{cov}(\underline{x}_i, \underline{x}_j) = \sigma_{ij} \quad ; \quad \text{cov}(\underline{x}_i, \underline{y}) = \sigma_{iy} \quad ; \quad \text{var}(\underline{x}_i) = \sigma_i^2 \\ \text{var}(\underline{v}^{p(i)}) = \sigma_{vi}^2 \quad \text{und} \quad \text{DET} = \text{Determinante der Kovarianzmatrix}$$

1) Die Bezeichnung 'Theorem' stellt keine Wertung der gefundenen Aussage dar, sondern wird lediglich aus Gründen der Kontinuität beibehalten

$$E(\hat{\gamma}_1, \hat{\gamma}_2^{(1)}, \hat{\gamma}_2^{(2)}) = \begin{bmatrix} \sigma_1^2 & a_1 \sigma_{12} & a_2 \sigma_{12} \\ a_1 \sigma_{12} & a_1^2 \sigma_{22}^2 + \sigma_{v1}^2 & a_1 a_2 \sigma_{22}^2 \\ a_2 \sigma_{12} & a_1 a_2 \sigma_{22}^2 & a_2^2 \sigma_{22}^2 + \sigma_{v2}^2 \end{bmatrix} \begin{bmatrix} \sigma_{1y} \\ a_1 \sigma_{2y} \\ a_2 \sigma_{2y} \end{bmatrix}$$

$$= \frac{1}{\text{DET}} \begin{bmatrix} \sigma_{1y} (\sigma_{v1}^2 \sigma_{v2}^2 + \sigma_{22}^2 (a_1^2 \sigma_{v2}^2 + a_2^2 \sigma_{v1}^2)) - \sigma_{2y} \sigma_{12} (a_1^2 \sigma_{v2}^2 + a_2^2 \sigma_{v1}^2) \\ a_1 \sigma_{v2}^2 (\sigma_1^2 \sigma_{2y} - \sigma_{12} \sigma_{1y}) \\ a_2 \sigma_{v1}^2 (\sigma_1^2 \sigma_{2y} - \sigma_{12} \sigma_{1y}) \end{bmatrix}$$

Man sieht also, daß in diesem Modell die Proxy-Variablen $p_2^{(1)}$ und $p_2^{(2)}$ zu Schätzungen der Regressionskoeffizienten führen, die durch die Gewichte $a_1 \sigma_{v2}^2$ bzw. $a_2 \sigma_{v1}^2$ gekennzeichnet sind. Wenn aber, wie in dem 'Theorem' 3 postuliert, eine der beiden Proxies die Variable x_1 selbst ist, also $\sigma_{v(x_1)}^2 = 0$ ist, erhält die 'echte' Proxy-Variable den Schätzkoeffizienten Null.

Eine Verallgemeinerung des kleinen Modells zur Aussage des 'Theorems' 3 mit $\sigma_{v(x_1)}^2 = 0$ könnte formal über die konsequente Anwendung der Adjunktenregel zur Invertierung von Matrizen erfolgen. Der Verfasser hält aber die folgende Herleitung für anschaulicher:

Da der OLS-Schätzer die Größe $\sum (y_t - \hat{y}_t)^2$ minimiert, wobei ja $\hat{y}$ die Orthogonalprojektion von y auf die Ebene der Regressoren darstellt mit dem Orthogonalabstand $\hat{v}$, würde jede Ebene der Regressoren mit $\lambda_1 x_1 + \lambda_2 p_1 = (\lambda_1 + \lambda_2) x_1 + \lambda_2 v^p$ mit einem $\lambda_2 \neq 0$ wegen $E(v' v^p) = 0$ diesen Abstand vergrößern.

Somit liefert 'Theorem' 3 in Verbindung mit Theorem 1 den für praktische 'Versuchsspezifikationen' vielleicht interessanten Hinweis, wann eine Proxy durch eine 'echte' Variable ersetzt wurde: Der geschätzte Regressionskoeffizient der Proxy wird zu Null und s^2 wird kleiner. Ein numerisches Beispiel ist das Modell des Anhangs auf Seite A. I., 4. Spezifikation.

§ 5. Seemingly unrelated regression - Systeme

Stellt man eine Hierarchie der Interdependenz von Regressionsgleichungssystemen auf, so kann man folgende Rangordnung anwenden:

1) Eingleichungssysteme : $\underline{y} = \underline{X}\underline{\gamma} + \underline{v}$

Eine solche Gleichung steht ohne Restriktionen durch andere Regressionsbeziehungen; auf die Schätzprobleme wurde schon in den vorgehenden Kapiteln eingegangen.

2) Seemingly unrelated regression - Systeme (SUR-Systeme):

$$\underline{y}_i = \underline{X}_i \underline{\gamma}_i + \underline{v}_i \quad (i=1, \dots, G)$$

Hier liegt ein ganzer Satz von Regressionsgleichungen vor, die jedoch jeweils nur eine abhängige Variable beinhalten. Dieses Modell und seine Schätzer werden im folgenden Abschnitt dieses Kapitels behandelt.

3) Rekursive Systeme (RS)

$$\underline{y}_i = \underline{Y} \underline{\beta}_i + \underline{X} \underline{\gamma}_i + \underline{u}_i \quad (i=1, \dots, G)$$

Hier liegt ebenfalls ein System von G Regressionsgleichungen vor, bei dem allerdings auch abhängige Variablen als Regressoren auftreten können. Durch Umnummerierung der einzelnen Gleichungen muß erreicht werden können, daß für die Elemente jedes Vektors $\underline{\beta}_i$ gilt: $\beta_{ij} = 0$ für $i \leq j$. Außerdem wird gefordert $E(\underline{u}_i \underline{u}_j) = 0$ für $i \neq j$. Dieses Modell, seine Schätzprobleme und Verallgemeinerungen der Modellannahmen werden im Kapitel 8 behandelt.

4) Interdependente Systeme (IS) :

$$\underline{YB} = \underline{X}\underline{\Gamma} + \underline{U} \quad 1)$$

Hier entfallen die Restriktionen des rekursiven Systems, dadurch werden zur Schätzbarkeit (Identifikation) zusätzliche Restriktionen des Systems nötig; deren Behandlung

1) Die verwendete Nomenklatur wird in Kapitel 7 noch ausführlich erklärt.

wird ab Kapitel 7 erfolgen.

Nach den Ausführungen des vergangenen Kapitels sind zusätzliche Informationen nötig, um konsistente Lösungsansätze für den Fall fälschlich ausgeschlossener Regressoren und für das EIV-Modell zu gewinnen. Der folgende Abschnitt wird untersuchen, ob die bei SUR-Systemen verwendeten Schätzer diesem Anspruch genügen.

5.1. Die Koeffizientenschätzung im SUR-System nach Zellner

Das bei der Schätzung von SUR-Systemen wohl am weitesten verbreitete Schätzverfahren geht auf Zellner ¹⁾ zurück und verwendet folgende Modellannahmen:

Gegeben sei ein Satz von Endogenen $\underline{y}_i$ ($i=1, \dots, G$), die jeweils durch eine Regressionsbeziehung

$$\underline{y}_i = X_i \underline{\gamma}_i + \underline{v}_i$$

mit den OLS-Modellannahmen des Kapitels 1 erklärt werden.

Das Gesamtsystem lautet damit

$$\begin{bmatrix} \underline{y}_1 \\ \vdots \\ \underline{y}_i \\ \vdots \\ \underline{y}_G \end{bmatrix} = \begin{bmatrix} X_1 & & & \emptyset \\ & \ddots & & \\ \emptyset & & X_i & \\ & & & \ddots \\ & & & & X_G \end{bmatrix} \begin{bmatrix} \underline{\gamma}_1 \\ \vdots \\ \underline{\gamma}_i \\ \vdots \\ \underline{\gamma}_G \end{bmatrix} + \begin{bmatrix} \underline{v}_1 \\ \vdots \\ \underline{v}_i \\ \vdots \\ \underline{v}_G \end{bmatrix} \quad \text{oder} \quad Y = X\Gamma + V$$

Dabei sind die $\underline{y}_i$ jeweils $T \cdot 1$ -Vektoren, die X_i jeweils $T \cdot K_i$ -Matrizen, die $\underline{\gamma}_i$ jeweils $K_i \cdot 1$ -Vektoren und die $\underline{v}_i$ jeweils $T \cdot 1$ -Vektoren. Weiter seien die X_i nicht für alle i ($i=1, \dots, G$) identische Matrizen, die Regressoren jeder Gleichung seien also in mindestens einer Spalte von X_i verschieden.

Mit den Annahmen

$$E(\underline{v}_i) = \emptyset \quad ; \quad E(\underline{v}_i \underline{v}_i') = \sigma_{ii} \cdot I \quad ; \quad E(X' \underline{v}_i) = \emptyset \quad \text{und} \quad E(\underline{v}_i \underline{v}_j') = \sigma_{ij} \cdot I \quad ^2)$$

1) A. Zellner (1962)

2) Also weder Autokorrelation noch Heteroskedastizität innerhalb und zwischen den Gleichungen

gilt für die Varianz-Kovarianz-Matrix des Gesamtsystems

$$E(VV') = \begin{bmatrix} \sigma_{11} I & \dots & \sigma_{1G} I \\ \vdots & & \vdots \\ \sigma_{G1} I & \dots & \sigma_{GG} I \end{bmatrix} = \Sigma \otimes I \quad 1)$$

Der Rang von Σ sei G , d. h. die y_i seien nicht linear abhängig. Auf dieses System nun wendet Zellner den Aitken-Schätzer (GLS = Generalized Least Squares) an nach

$$\hat{\beta} = (X'(\Sigma^{-1} \otimes I)X)^{-1}X'(\Sigma^{-1} \otimes I)Y$$

mit der Varianz-Kovarianz-Matrix

$$\text{var}(\hat{\beta}) = (X'(\Sigma^{-1} \otimes I)X)^{-1}$$

Für identische X_i fallen die Schätzungen nach Zellner und die OLS-Schätzungen der Einzelgleichungen zusammen.²⁾

Zur Berechnung von $\hat{\beta}$ muß allerdings zuerst die ja unbekannte Matrix Σ geschätzt werden. Hierzu werden verschiedene Vorschläge gemacht³⁾, die aber alle mit einer noch später gesondert zu behandelnden Ausnahme auf einem zweistufigen Schätzverfahren beruhen:

In der ersten Stufe werden nach OLS-Schätzungen aller Gleichungen ohne irgendwelche Restriktionen die Residuen dieser Regressionsschätzungen bestimmt nach $\hat{u}_i^{(1)} = y_i - X(X'X)^{-1}X'y_i$. Hieraus werden die $\hat{\sigma}_{ij}$ bestimmt nach $\hat{\sigma}_{ij} = \hat{u}_i^{(1)} \hat{u}_j^{(1)} / F_{ij}$, wobei die F_{ij} die "Freiheitsgrade" der Schätzung sind.

Das hier auftauchende Problem - K_i und K_j als Anzahl der in den verschiedenen Gleichungen auftretenden Regressoren müssen ja nicht übereinstimmen - führt zu unterschiedlichen Vorschlägen. Johnston⁴⁾ schlägt ohne Begründung $F_{ij} = \sqrt{(T-K_i)(T-K_j)}$ vor, Zellner⁵⁾ und Rao⁶⁾ benutzen $F_{ij} = T$, Theil⁷⁾ benutzt

1) $\otimes$ ist das Zeichen für das Kronecker-Produkt

2) Vgl. z. B. Theil (1971) S. 309 f.

3) Telser (1964); Kmenta & Gilbert (1968)

4) Johnston (1972) A. 240

5) Zellner (1963) S. 980

6) Rao (1974)

7) Theil (1971) S. 321

neben $F_{ij} = T$ auch $F_{ij} = T - K_i - K_j + \bar{r}_{ij}^2 \cdot \min(K_i, K_j)$, wobei $\bar{r}_{ij}^2$ das Quadrat des kanonischen Korrelationskoeffizienten zwischen X_i und X_j ist. Als einziger weist Theil ¹⁾ darauf hin, daß man selbst mit einer unverzerrten Schätzung $\hat{\Sigma}$ mit $\hat{\Sigma}^{-1}$ keine unverzerrte Schätzung für Σ^{-1} erhält. Für alle obigen Vorschläge muß F_{ij} selbstverständlich positiv sein, was entsprechende Forderungen an den Stichprobenumfang T stellt.

In der zweiten Stufe wird dann der Aitken - Schätzer angewandt, wobei die Matrix Σ^{-1} durch die Schätzung $\hat{\Sigma}^{-1}$ der ersten Stufe ersetzt wird. Beläßt man es bei dieser zweiten Stufe, so liegt der "Zweistufige Zellner - Schätzer" vor, bei dem allerdings stört, daß die Varianz-Kovarianz-Matrix $\Sigma^{(2)}$ des Gesamtsystems im allgemeinen nicht mit derjenigen übereinstimmt, die zur Bestimmung der Regressionskoeffizienten benutzt wurde: Durch die Veränderung der Regressionskoeffizienten der zweiten Stufe gegenüber denen der ersten (OLS -) Stufe ändern sich die $\hat{v}^{(2)}$ gegenüber den $\hat{v}^{(1)}$ und damit auch $\Sigma^{(2)}$ gegenüber $\Sigma^{(1)}$. Wie Zellner schon in seiner ersten Arbeit andeutet ²⁾, kann man nun in weiteren Schritten iterativ vorgehen und für jede r -te Schätzung der Regressionskoeffizienten die Inverse der Schätzung von $\hat{\Sigma}^{(r-1)}$ verwenden.

Über Abbruchkriterien und Konvergenzverhalten dieses Verfahrens ist in der Literatur nichts zu finden ; eigene Simulationen des Verfassers zeigten allerdings, daß sich beifast allen untersuchten simulierten Modellen schon nach wenigen Iterationen kein Element aus $\hat{\Sigma}^{(r)}$ um mehr als 10^{-5} änderte ³⁾, das Verfahren also 'gutmütig' zu sein scheint.

Auf eines muß allerdings hingewiesen werden: Wie Kmenta und

-
- 1) Theil (1971) S. 322. Er weist außerdem darauf hin, daß das Gauß-Markow-Theorem nicht mehr anwendbar ist, da der Schätzer nichtlinear ist wegen seiner Abhängigkeit von Σ^{-1} (S.324)
 2) Zellner (1962) S. 363 Fußnote 13
 3) Kmenta & Gilbert (1968) S. 1186 brachen erst bei Veränderungen unter 10^{-9} ab.

Gilbert ¹⁾ zeigen, sind sowohl der zweistufige wie der iterative Zellner-Schätzer konsistent und asymptotisch effizient. Wegen der Gleichheit der asymptotischen Eigenschaften kann sich eine Überlegenheit des iterativen gegenüber dem zweistufigen Schätzer nur in kleinen Stichproben zeigen. Eine solche Überlegenheit konnte in ihren Simulationen aber nicht gefunden werden.

Unter der Annahme symmetrisch verteilter, stetiger Störgrößen v ist der Schätzer von Zellner unverzerrt ²⁾ und, mit der zusätzlichen Annahme $X_i'X_j = \emptyset$, asymptotisch effizient gegenüber den einfachen OLS-Schätzungen ³⁾. Für endliche Meßreihen zeigt Zellner an einem Modell mit $G = 2$ tabellarisch auf, wie die Effizienz seines Schätzers gegenüber einfachem OLS mit wachsendem ρ_{ij} zunimmt ⁴⁾.

An dieser Stelle ist, vor weiteren Ausführungen, die Bedeutung der "scheinbar nicht zusammenhängenden Regressionen" (SUR) zu untersuchen, wobei vor allem den möglichen Ursachen für das Auftreten eines $\sigma_{ij} \neq 0$ nachzugehen ist.

Grundannahme des SUR-Schätzverfahrens ist eine richtige Spezifikation aller Gleichungen des Gesamtsystems, so daß sie mit OLS unverzerrt und konsistent geschätzt werden können. Wenn man aber nach diesen OLS-Schätzungen eine wechselseitige Korrelation der Residuen der Einzelgleichungen feststellt, sind diese nur noch "scheinbar" unzusammenhängend. Diese 'mutual correlation' geht dann als zusätzliche Information in den Aitken-Ansatz des Zellner-Schätzers ein und kann damit die Effizienz des Schätzers verbessern: Es liegt nun ein 'full

1) Kmenta & Gilbert (1968)

2) N. C. Kakwani (1967). Seine Aussage bezieht sich lediglich auf den zweistufigen Zellner-Schätzer.

3) Zellner & Huang (1962)

4) Zellner (1963); ρ ist der Korrelationskoeffizient der Gesamtheit.

information' - Schätzverfahren vor, das die Gesamtinformation des Gleichungssystems berücksichtigt ¹⁾. Dieser Effizienzgewinn wird aber, wie das ja auch von den 'full information'-Schätzern bei interdependenten Systemen bekannt ist ²⁾, durch eine erhöhte Empfindlichkeit gegen ein fehlspezifiziertes System bezahlt. Hier setzt die Kritik z. B. von Maddala ³⁾ an, der nach der Herkunft der 'mutual correlation' fragt. Denn wenn diese nur das Produkt von im endlichen Fall auftretenden Abweichungen der geschätzten $\hat{v}$ von den wahren v sind, verschwindet diese Korrelation für $T \rightarrow \infty$, SUR - und OLS - Schätzer fallen zusammen. Da aber die Residuen v ja die Summation ausgeschlossener Regressoren darstellen ⁴⁾, liegt der Verdacht nahe, daß die 'mutual correlation' zwischen den Einzelgleichungen durch das Ausschließen des- oder derselben Regressoren aus der i-ten und j-ten Gleichung verursacht wurde. Und gerade für diesen Fall zeigt sich, daß der zweistufige SUR-Schätzer keine Überlegenheit gegenüber einfachem OLS aufweist; Maddala spricht sogar von "Versagen" ⁵⁾. Dieses Ergebnis ist enttäuschend, aber nicht überraschend, da ja der im SUR - Ansatz verwendete Aitken - Schätzer von falschen Voraussetzungen ausgehend nicht die Ursachen behandelt, sondern mit den Auswirkungen rechnet. Zur Verdeutlichung soll wieder ein kleines Modell dienen: ⁶⁾:

Gegeben sei das wahre Modell

$$Y_1 = X_1 \gamma_{11} + X_3 \gamma_{13} + v_1; \text{ also } X_1 = (X_1; X_3)$$

$$Y_2 = X_2 \gamma_{22} + X_3 \gamma_{23} + v_2; \text{ also } X_2 = (X_2; X_3)$$

-
- 1) Dies wird z. B. von P. Schmidt (1978) schön veranschaulicht
 2) Vgl. z. B. Schneeweiß (1974) S. 318
 3) Maddala (1977) S. 321
 4) Vgl. S. 3 und S. 5 Fußnote 2)
 5) Er bezieht sich dabei auf Rao (1974). Vgl. jedoch Kapitel 6!
 6) Ein numerisches Beispiel ist im Anhang Seite A.IV.; zu beachten ist, daß das Beispiel dem Fall 4 bei Rao entspricht, bei dem der SUR-Schätzer nicht stärker verzerrt sein muß!

mit den üblichen Regressionsannahmen $E(X_i' v_j) = 0$ für i und $j = 1; 2$. Außerdem seien $\underline{x}_1$, $\underline{x}_2$ und $\underline{x}_3$ linear unabhängig. Wird jetzt fälschlicherweise die Variable $\underline{x}_3$ aus beiden Regressionsansätzen ausgeschlossen, so ergibt der OLS - Regressionsansatz für

$$\begin{aligned}\hat{v}_1 &= y_1 - x_1(x_1'x_1)^{-1}x_1'y_1 = \hat{v}_1 + (I - x_1(x_1'x_1)^{-1}x_1')x_3\gamma_{13} \quad \text{und} \\ \hat{v}_2 &= y_2 - x_2(x_2'x_2)^{-1}x_2'y_2 = \hat{v}_2 + (I - x_2(x_2'x_2)^{-1}x_2')x_3\gamma_{23},\end{aligned}$$

wobei die $\hat{v}_i$ die geschätzten OLS - Residuen bei vollständiger Spezifikation sind. Nimmt man an, daß $E(\hat{v}_1\hat{v}_2) = 0$ ist, so ist ¹⁾

$$\begin{aligned}E(\hat{v}_1\hat{v}_2) &= E(((I - x_1(x_1'x_1)^{-1}x_1')x_3\gamma_{13})'((I - x_2(x_2'x_2)^{-1}x_2')x_3\gamma_{23})) \\ &= \gamma_{13}\gamma_{23}E(\hat{e}_1\hat{e}_2),\end{aligned}$$

wobei die $\hat{e}_i$ die Residuen einer Hilfsregression von $\underline{x}_3$ auf $\underline{x}_1$ bzw. $\underline{x}_2$ sind nach

$$\hat{e}_1 = \underline{x}_3 - \underline{x}_1(\underline{x}_1'\underline{x}_1)^{-1}\underline{x}_1'\underline{x}_3 = \underline{x}_3 - \hat{\underline{x}}_3(1)$$

Diese Beziehung wird in Abbildung 1a im Testraum dargestellt:

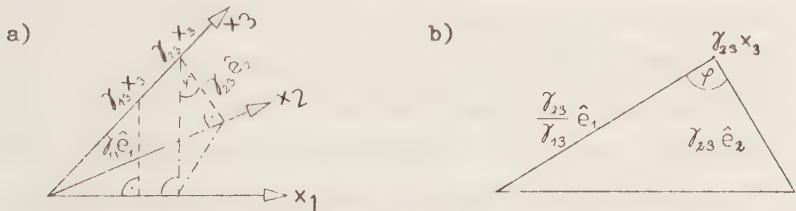


Abb. 1

Wie man aus der Darstellungsweise im Testraum weiß, ist nur für $\cos \varphi = 0$ die Kovarianz zwischen $\hat{e}_1$ und $\hat{e}_2 = 0$ (Abb. 1b). Für $\cos \varphi = 0$ oder $\varphi = 90^\circ$ muß aber nach dem Satz des Pythagoras gelten

$$\|\hat{e}_1 \cdot \gamma_{23}/\gamma_{13}\|^2 + \|\hat{e}_2 \cdot \gamma_{23}\|^2 = \|\hat{\underline{x}}_3(1) \cdot \gamma_{23}/\gamma_{13} - \hat{\underline{x}}_3(2) \cdot \gamma_{23}\|^2,$$

wobei z. B. $\|\hat{e}_2\|^2$ das Quadrat der Länge von $\hat{e}_2$ bzw. seine Variation darstellt. Auch diese Beziehung ist in Abb. 1b dargestellt. Für jede Lage von $\underline{x}_3$ also, die nicht auf dieser durch obige Beziehung definierten Kegeloberfläche liegt, ist

1) Vgl. Kapitel 2.1.2. S. 18

die Kovarianz zwischen $\hat{e}_1$ und $\hat{e}_2$ ungleich Null. Das heißt, daß $\hat{e}_1$ und $\hat{e}_2$ nur im Ausnahmefall orthogonal sein werden. Oder aber man kann, in Umkehrung der Betrachtungsweise, aus einer Kovarianz $\sigma_{ij} \neq 0$ auf eine gemeinsame (nämlich bezüglich $\underline{x}_3$) Fehlspezifikation des Typs II schließen. Damit sind die beiden alternativen Schätz - "Strategien" gegenübergestellt: Entweder betrachtet man die Kovarianz zwischen den Einzelgleichungen eines SUR - Systems als ein Mittel, die Effizienz der ohnehin schon unverzerrten OLS - Schätzungen zu verbessern. Oder man geht davon aus, daß diese OLS - Schätzungen und die aus ihnen resultierenden Residuen auf einen oder mehrere ausgeschlossene Regressoren schließen lassen, weil man davon ausgeht, daß Zufallsgrößen nur dann korreliert sind, wenn sie gemeinsame Elemente besitzen. Wenn man weiterhin davon ausgeht, daß der Ausschluß von Regressoren im allgemeinen eine Inkonsistenz der OLS - Schätzungen nach sich zieht, so ist das Ziel einer 'full information' - Schätzung der zweiten Schätzalternative die Beseitigung oder zumindest Verkleinerung der Inkonsistenz. Diese zweite Alternative wird im folgenden Kapitel eingeführt.

§6. Alternativer SURE - Ansatz

=====

Wie in den vorhergehenden Kapiteln gezeigt wurde, stellen sowohl Unterspezifikation wie auch die Verwendung fehlerbehafteter Regressoren bei der Anwendung von OLS eine ernstzunehmende Ursache für inkonsistente Schätzungen der Regressionskoeffizienten dar. Dies war auf einen abgestuften Mangel an Informationen zurückgeführt worden: Im Fall der Unterspezifikation weiß man gar nichts über die fehlenden Variablen, im Fall fehlerbehafteter Regressoren hat man wenigstens eine Proxy-Variable für die fehlende Variable. Bei SUR - Systemen erhält man aber zusätzliche Informationen, die in Verbindung mit a - priori - Annahmen eine Verbesserung der Schätzansätze erlauben.

Zur Illustration diene wieder das kleine SUR - Modell des vorangegangenen Kapitels:

$$Y_1 = X_1 \gamma_{11} + 0 + X_3 \gamma_{13} + v_1 = X_1 \gamma_1 + v_1 \quad (1.1)$$

$$Y_2 = 0 + X_2 \gamma_{22} + X_3 \gamma_{23} + v_2 = X_2 \gamma_2 + v_2 \quad (1.2)$$

mit den dort angeführten Annahmen

$$\text{rg}(X_1) = \text{rg}(X_2) = 2 ; E(X_i' v_j) = 0 \text{ für } i, j=1,2 ,$$

die jetzt um die folgenden Annahmen erweitert werden

$$\text{rg}(X_1; X_2) = 2 \quad (2)$$

$$E(v_1' v_2) = 0 \quad (3)$$

Dieses Modell kann man umformulieren mit

$$X_3 = (1/\gamma_{13})(Y_1 - v_1 - X_1 \gamma_{11}) = (1/\gamma_{23})(Y_2 - v_2 - X_2 \gamma_{22}) \quad (4)$$

zu

$$Y_1 = X_1 \gamma_{11} + (\gamma_{13}/\gamma_{23})(Y_2 - v_2 - X_2 \gamma_{22}) + v_1 \quad (5.1)$$

$$Y_2 = X_2 \gamma_{22} + (\gamma_{23}/\gamma_{13})(Y_1 - v_1 - X_1 \gamma_{11}) + v_2 \quad (5.2)$$

Diese beiden Gleichungen sind linear abhängig, d. h. die zweite Gleichung ist das $(\gamma_{23}/\gamma_{13})$ -fache der ersten Gleichung.

6.1. SURE - Ansatz für das 'EIV' - Modell

Für dieses kleine Modell soll nun der Fall "Fehler im Regressor" untersucht werden. Fehlerbehaftet sei die Variable $\underline{x}_3$ nach $\underline{x}_3^+ = \underline{x}_3 + \underline{v}^+ = \underline{p}_3$ mit den "klassischen" Annahmen

$$\text{plim}(\frac{1}{T} \underline{x}_1' \underline{v}^+) = \text{plim}(\frac{1}{T} \underline{x}_2' \underline{v}^+) = \text{plim}(\frac{1}{T} \underline{x}_3' \underline{v}^+) = 0 \quad (7)$$

$$\text{plim}(\frac{1}{T} \underline{v}_1' \underline{v}^+) = \text{plim}(\frac{1}{T} \underline{v}_2' \underline{v}^+) = 0 \quad (8)$$

Dann gilt: $\underline{y}_2$ ist eine Instrumentalvariable (IV) für $\underline{x}_3^+ = \underline{p}_3$, denn sie erfüllt die an eine IV gestellten Bedingungen ¹⁾

$$\text{plim}(\frac{1}{T} \underline{y}_2' \underline{v}_1) = \text{plim}(\frac{1}{T} \underline{y}_2' \underline{v}^+) = 0 \quad (9)$$

$$\text{plim}(\frac{1}{T} \underline{y}_2' \underline{x}_3) = \sigma(\underline{y}_2 \underline{x}_3) \text{ existiert und ist nicht } 0 \quad (10)$$

$$\text{plim}(\frac{1}{T} \underline{y}_2' \underline{y}_1) = \sigma(\underline{y}_2 \underline{y}_1) \text{ existiert und ist nicht } 0 \quad (11)$$

Damit ist ein IV - Ansatz möglich nach

$$\hat{\underline{z}}_1 = (\underline{Z}' \underline{X}^+)^{-1} \underline{Z}' \underline{y}_1 \quad (12)$$

wobei $\underline{X}^+ = (\underline{x}_1; \underline{x}_3^+)$ und $\underline{Z} = (\underline{x}_1; \underline{y}_2)$.

Denn schreibt man

$$\underline{Z}' \underline{y}_1 = \underline{Z}' \underline{x}_1 \underline{\gamma}_1 + \underline{Z}' \underline{v}_1, \quad (13)$$

so ist der Wahrscheinlichkeitslimes $\text{plim}(\frac{1}{T} \underline{Z}' \underline{v}_1) = 0$

und damit $\hat{\underline{z}}_1$ ein konsistenter Schätzer für $\underline{\gamma}_1$ durch Auflösen nach $\underline{\gamma}_1$, was man sofort sieht nach

$$\begin{aligned} \hat{\underline{z}}_1 &= \underline{\gamma}_1 + (\underline{Z}' \underline{X}^+)^{-1} \underline{Z}' \underline{v}_1 \quad \text{und} \\ \text{plim}(\hat{\underline{z}}_1) &= \underline{\gamma}_1 + \text{plim}(\frac{1}{T} \underline{Z}' \underline{X}^+)^{-1} \cdot \text{plim}(\frac{1}{T} \underline{Z}' \underline{v}_1) \\ &= \underline{\gamma}_1 + \text{plim}(\frac{1}{T} \underline{Z}' \underline{X}^+)^{-1} \cdot 0 = \underline{\gamma}_1 \end{aligned} \quad (14)$$

Die asymptotische Varianz-Kovarianz-Matrix von $\hat{\underline{z}}$ läßt sich dann darstellen als ²⁾

$$\text{asy var}(\hat{\underline{z}}_1) = \frac{1}{T} \sigma_{\underline{v}_1}^2 \sum \underline{Z} \underline{Z}' \sum \underline{X}^+ \underline{X}^+{}^{-1} \quad (15)$$

1) Vgl. z. B. Johnston (1972) S. 279 f., besonders S. 283

2) Vgl. z. B. Johnston (1972) S. 283

die konsistent abzuschätzen ist nach ¹⁾

$$\text{asy var}(\hat{\beta}) = s^2 (Z'X^+)^{-1}(Z'Z)(X^+Z)^{-1} \quad (16)$$

mit

$$s^2 = [(Z_1'Z_1 - (X_1'\hat{\beta}_1)'Z_1)] \cdot \frac{1}{Fr} \quad 2) \quad (17)$$

Dabei kann man den Nenner "Fr" zu T wählen und erhält

damit zumindest einen konsistenten Schätzer für $\sigma_{v_1}^2$.

In Analogie zum OLS - Schätzer kann, wie z. B. von Johnston ³⁾,

Fr = T - K gewählt werden; allerdings muß man sich dabei vor der Vorstellung hüten, damit einen unverzerrten Schätzer für

$\sigma_{v_1}^2$ zu erhalten: Da für $\hat{\beta}$ nur die Konsistenz mit $T \rightarrow \infty$ bewiesen wurde, kann die Eigenschaft einer von $\hat{\beta}$ abhängigen Größe wie s^2 auch nur für $T \rightarrow \infty$ angegeben werden ⁴⁾.

Aber ebenso, wie z. B. bei der Methode der zweistufigen

kleinsten Quadrate 2SLS die (asymptotische) Effizienz der

Koeffizientenschätzung durch Verwendung nicht nur der Mindest-

anzahl an nötigen, sondern durch die Verwendung aller möglichen

zur Verfügung stehenden Instrumentalvariablen gesteigert werden

kann, ist dies bei dem vorliegenden Fall durch gleichartiges

Vorgehen möglich ⁵⁾:

Man bestimme in der ersten Stufe

$$\hat{x}_3 = (x_2, y_2)[(x_2, y_2)'(x_2, y_2)]^{-1}(x_2, y_2)'x_3 \quad (18)$$

und verwende in der zweiten Stufe $\hat{x}_3$ als IV für x_3^+ nach

$$\hat{\beta}_1 = (Z'X^+)^{-1}Z'y_1, \text{ jetzt aber mit}$$

$$Z = (x_1, \hat{x}_3) \text{ und } X^+ = (x_1, x_3^+) \quad (19)$$

1) Vgl. z. B. Johnston (1972) S. 283

2) Vgl. Kapitel 3.5.1. S. 29. Johnston gibt hier keinen Schätzer s^2 an; der von ihm allgemein für den IV-Ansatz auf S. 280 angegebene Schätzer ist hier nicht gültig.

3) Johnston (1972) S. 280

4) Das gleiche Problem wird noch einmal in Kapitel 7 auftauchen

5) Vgl. auch Dhrymes (1970) S. 298

Denn auch $\hat{x}_3$ erfüllt die notwendigen Bedingungen, die an eine Instrumentalvariable gestellt werden ¹⁾ :

$$\text{plim}(\frac{1}{T} \hat{x}_3' y_1) = \text{plim}(\frac{1}{T} \hat{x}_3' v^+) = 0 \quad (20)$$

$$\text{plim}(\frac{1}{T} \hat{x}_3' x_1) = \sigma_{x_1 \hat{x}_3} \text{ existiert und ist nicht } 0 \quad (21)$$

Und für $x_2' x_3^+ = x_2' x_3 \neq 0$ ist die Korrelation zwischen $\hat{x}_3$ und x_3^+ höher als die zwischen x_3^+ und y_2 allein; damit aber ist die in der $x_2 y_2$ - Ebene liegende IV $\hat{x}_3$ eine bessere IV als y_2 allein ²⁾.

In Fortsetzung der Analogie des obigen Schätzers zur Methode der zweistufigen kleinsten Quadrate kann nun auch der Fall betrachtet werden, daß x_3 in mehr als zwei Gleichungen eines SURE - Systems auftritt. Da dann die Zahl der vorhandenen Instrumentalvariablen die Zahl der zur Identifikation nötigen überschreitet, ist bei dem einfachen einstufigen Instrumentalvariablenansatz, bei dem dann z. B. y_2 oder y_3 direkt als Instrumentalvariable dient, eine Auswahl unter diesen beiden Instrumentalvariablen zu treffen. Im zweistufigen IV - Ansatz nach (18) und (19) dagegen werden alle im System vorhandenen Instrumentalvariablen verwendet, also z. B. y_2 , y_3 und die in der 2. und 3. Gleichung spezifizierten Exogenen. Damit ist der zweistufige Ansatz nach (18) und (19) eine Verallgemeinerung des z. B. von Maddala ³⁾ vorgestellten einfachen Instrumentalvariablenansatzes. Der entscheidende Unterschied zu 2SLS wie überhaupt zu allen Instrumentalvariablenschätzern für interdependente Strukturgleichungen liegt in der Definition der Residualvarianz nach (17): s^2 kann hier nicht nach $s^2 = \sum_t (y_{1t} - x_t' \hat{\gamma}_1)^2$ bestimmt werden, wie es die Analogie fordern würde ⁴⁾.

1) Vgl. z. B. Schönfeld (1971) S. 239

2) Vgl. z. B. Wonnacott & Wonnacott (1970) S. 358 f. und vor allem Dhrymes (1970) S. 298

3) Maddala (1977) S. 301

4) Dies ist überhaupt der wesentliche Unterschied zwischen interdependenten Systemen und EIV - Modellen!

Verallgemeinert man den Ansatz (19) auf ein größeres System, das mehr als eine Proxy-Variable beinhaltet, so erhält man

$$\begin{bmatrix} \underline{y}_1 \\ \vdots \\ \underline{y}_i \\ \vdots \\ \underline{y}_G \end{bmatrix} = \begin{bmatrix} X_1 & & & & \\ & \ddots & & & \\ & & X_i & & \\ \emptyset & & & \ddots & \\ & & & & X_G \end{bmatrix} \begin{bmatrix} \underline{z}_1 \\ \vdots \\ \underline{z}_i \\ \vdots \\ \underline{z}_G \end{bmatrix} + \begin{bmatrix} \underline{v}_1 \\ \vdots \\ \underline{v}_i \\ \vdots \\ \underline{v}_G \end{bmatrix} \quad (22)$$

mit den Annahmen

$$\text{plim}(\frac{1}{T} X_i' \underline{v}_i) = \text{plim}(\frac{1}{T} \underline{v}_i' \underline{v}_i) = 0 \quad \text{für } i', i, j=1, \dots, G; i \neq j \quad (23)$$

$$E(\underline{v}_{it}) = 0 \quad \text{und} \quad E(\underline{v}_i \underline{v}_i') = \sigma_i^2 \cdot I \quad \text{für } i=1, \dots, G; t=1, \dots, T \quad (24)$$

Weiterhin seien in einer Matrix $X^+ = P = (P_1; P_2; \dots; P_{\lambda}; \dots; P_l)$ l Proxyvariablen in mindestens $l' = l+1$ Gleichungen enthalten, die o. B. d. A. die l'-ersten seien. Dabei soll gelten

$$P_{\lambda} = \underline{x}_{\lambda} + \underline{v}_{\lambda}^p \quad (25)$$

$$\text{plim}(\frac{1}{T} \underline{v}_{\lambda'}^p \underline{v}_{\lambda}^p) = 0 \quad \text{für } \lambda', \lambda=1, \dots, l \text{ und } \lambda \neq \lambda' \quad (26)$$

$$\text{plim}(\frac{1}{T} \underline{v}_{\lambda'}^p \underline{v}_i) = 0 \quad \text{für } i=1, \dots, l'; \lambda=1, \dots, l \quad (27)$$

$$\text{plim}(\frac{1}{T} X_i' \underline{v}_{\lambda}^p) = 0 \quad \text{für } i=1, \dots, l'; \lambda=1, \dots, l \quad (28)$$

Das heißt, daß für jede in einer Gleichung auftretende Proxy mindestens eine andere Gleichung vorhanden sein muß, die diese Proxy beinhaltet, wobei die Fehler in den Regressoren sowohl mit den Regressoren wie auch wechselseitig im Wahrscheinlichkeitslimites unkorreliert sein sollen ebenso wie die Fehler in den Gleichungen untereinander. An die X_i werden lediglich die Bedingung gestellt, daß die in ihnen enthaltenen $\underline{x}_{ij}$ ($j=1, \dots, K_i$) nicht kollinear sind, der Rang von X_i also K_i ist; die X_i können aber (im Gegensatz zu späteren Ansätzen) für alle Gleichungen identisch sein, das heißt $X_i = X$. Dann lautet der Schätzansatz in der ersten Stufe mit den Bezeichnungen

$$X_z = X_2 \cup X_3 \cup X_4 \cup \dots \cup X_l \quad (29)$$

$$Y_z = \underline{y}_2 \cup \underline{y}_3 \cup \underline{y}_4 \cup \dots \cup \underline{y}_l \quad (30)$$

d. h. X_z und Y_z sind die Vereinigungsmengen der exogenen und endogenen Variablen der 2-ten bis $l^{\wedge}$ -ten Gleichung,

$$\hat{p}_\lambda = (X_z, Y_z) [(X_z, Y_z)'(X_z, Y_z)]^{-1} (X_z, Y_z)' p_\lambda \quad (31)$$

Diese $\hat{p}$ werden dann in der zweiten Stufe als Instrumentalvariablen verwendet nach

$$\hat{z}_1 = (Z_1' X_1)^{-1} Z_1' y_1$$

wobei jetzt die Definitionen gelten

$$Z_1 = (\hat{p}_1, \hat{p}_2, \dots, \hat{p}_1, x_{1+1}, x_{1+2}, \dots, x_{K_1}) \quad (32)$$

und

$$X_1 = (p_1, p_2, \dots, p_1, x_{1+1}, x_{1+2}, \dots, x_{K_1}) \quad (33)$$

Die asymptotische Varianz-Kovarianz-Matrix von $\hat{z}_1$ ist wieder abzuschätzen nach

$$\text{asy var}(\hat{z}_1) = s_1^2 (Z_1' X_1)^{-1} (Z_1' Z_1) (X_1' X_1)^{-1}$$

mit

$$s_1^2 = (y_1' y_1 - (X_1' \hat{z}_1)' y_1) \cdot \frac{1}{Fr}$$

wobei Fr nach den Überlegungen der Seite 61 gewählt werden kann. Im numerischen Anhang Seite A.V. und A.VI. sind zwei kleine Beispiele vorgeführt mit jeweils einer Proxy-Variablen; während im Beispiel A.V. die X_1 für die beiden Gleichungen des Systems identisch sind, haben die beiden Gleichungen des Beispiels A.VI. lediglich die Proxy-Variable gemeinsam. Und gerade das zweite Beispiel zeigt auf Seite A.VI.2 den Anstieg der Effizienz durch das Wachsen der 't' - Werte, wenn für p_3 nicht nur x_2 , sondern (x_2, y_2) als Instrumentalvariable gewählt wird ¹⁾.

Zusammenfassend kann man die Zielsetzung des vorgeführten

1) Dies gilt für die erste Gleichung; analoges ist ebenso für die zweite Gleichung zu finden. Die hohe Effizienz der Schätzung bei Wahl nur von y_2 als IV für p_3 liegt an der Modellkonstruktion: γ_{13} und γ_{23} besitzen hohe Werte.

Schätzverfahrens durch die zugrundeliegende Modellvorstellung beschreiben. Es wird davon ausgegangen, daß in einem vollständig spezifizierten, d. h. nicht unter-spezifizierten ¹⁾ Modell eine Kovarianz zwischen den einzelnen Gleichungen des Modells im Wahrscheinlichkeits-limes nur dann Werte ungleich Null annehmen kann, wenn die Gleichungen des Modells denselben Spezifikationsfehler aufweisen. Dieser Spezifikationsfehler ist im behandelten Modell die in den Gleichungen gemeinsame Wahl des falschen Regressorentyps, d. h. der gemeinsamen Wahl einer oder mehrerer Proxy-Variablen als Regressoren: Die im Wahrscheinlichkeits-limes "unrelated regressions" werden also nur durch diesen gemeinsamen Spezifikationsfehler nicht verschwindende Kovarianzen aufweisen. Mit dieser Annahme und der zusätzlichen Voraussetzung, daß dem Ökonometriker bekannt ist, welche Variablen er als Proxy-Variablen eingeführt hat ²⁾, ist bei einer ausreichenden Anzahl identisch fehlspezifizierter Gleichungen ein konsistenter Schätzansatz zur Bestimmung der Regressionskoeffizienten möglich: Jede in einer Gleichung auftretende Proxy verlangt die Existenz mindestens einer weiteren Gleichung, die eben diese Proxy beinhaltet. Sofern der Ökonometriker an den Koeffizienten der als Instrumentalvariable dienenden Gleichung nicht interessiert ist, muß diese Gleichung auch nicht vollständig spezifiziert sein ³⁾. Diese Modellvorstellung weicht also deutlich von der im SUR-Schätzer nach Zellner zugrundegelegten ab, wo ja von einer bei korrekter Spezifikation des Modells nicht verschwindenden Kovarianz zwischen den Gleichungen

1) Vgl. hierzu die Ausführungen §2 auf S. 15

2) Dies ist bei der Vorstellung eines empirischen Modells meist zu erkennen an der Formulierung "die Variable p steht für die Variable $\underline{x}$ " oder ähnlichen Substitutionsbeschreibungen.

3) Vgl. z. B. das Beispiel im Anhang S. A.VI.2, wo schon $\underline{y}_2$ allein als IV dient; die Spezifikation der 2. Gleichung ist dabei unerheblich.

ausgegangen wird, die nicht zu einer Beseitigung einer Verzerrung, sondern zur Steigerung der Effizienz benutzt wird. Daß dieser Ansatz bei dem vorliegenden Modell "versagen" ¹⁾ muß, zeigt schon das Beispiel des Anhangs A.V. : Da hier $X_1 = X_2$ ist, ist der SUR-Schätzer mit dem OLS - Schätzer identisch ²⁾ und weist damit eine ebenso hohe Verzerrung auf wie dieser.

Die Annahme im Wahrscheinlichkeitslimes verschwindender Kovarianzen zwischen den Gleichungen ist ohne Zweifel eine starke Restriktion und muß deshalb von dem Ökonometriker aufgrund seines Modellverständnisses überprüft werden. Hierzu schreibt z. B. Fisher ³⁾ : "... if the disturbances are thought of as being the net effects of numerous independently acting small influences omitted from ... the equations, the assumption of uncorrelated disturbance terms for two equations is then the assumption that the omitted variables affecting the two equations have no large number of elements in common ⁴⁾ ". Da die Spezifikation der als Instrumentalvariable dienenden Gleichung oder Gleichungen bei dem vorliegenden Ansatz ja keine Rolle spielt, bedeutet diese Aussage für das vorliegende Modell, daß die zu schätzende Gleichung keine in y subsumierten, aus dem Schätzansatz ausgeschlossenen Einflußgrößen beinhalten darf, die auch die als IV dienenden anderen Endogenen beeinflussen. Genau diesen Fall aber wird der folgende Schätzansatz behandeln.

1) Dies ist eine Bestätigung der Feststellung von Maddala, wie sie auf S. 56 zitiert wurde.

2) Vgl. z. B. Johnston (1972) S. 240

3) Fisher (1966) S. 92

4) Vgl. hierzu auch die Ausführungen des Kapitels 1, vor allem S. 5 f.

6.2. Inkonsistenter SURE-Ansatz bei Unterspezifikation

Gegeben sei wieder das kleine Modell mit den Gleichungen (1.1) und (1.2), das neben den auf S. 59 aufgeführten üblichen Regressionsannahmen auch wieder die zusätzlichen Bedingungen (2) und vor allem (3) erfüllt.

Zusätzlich sei jetzt aber angenommen, daß weder die Variable $\underline{x}_3$ noch eine Proxy $\underline{p}_3$ gemessen werden kann; von dem Modell ist also lediglich bekannt, daß $\underline{y}_1$ neben $\underline{x}_1$ und $\underline{y}_2$ neben $\underline{x}_2$ noch beide von einer unbekannten Variable $\underline{x}_3$ beeinflusst werden. Wie schon mit der Umformulierung (4) gezeigt wurde, kann man also z. B. die erste Gleichung darstellen nach

$$\underline{y}_1 = \underline{x}_1 \gamma_{11} + (\gamma_{13}/\gamma_{23})(\underline{y}_2 - \underline{v}_2 - \underline{x}_2 \gamma_{22}) + \underline{v}_1 \quad (5.1)$$

Zur Unterscheidung des später einzuführenden Schätzverfahrens von dem auf Seite 52 als Ausnahme angedeuteten Verfahren soll nun vorläufig ein etwas umständlich wirkendes Schätzverfahren gezeigt werden: In einer ersten Stufe schätze man mit OLS

$$\hat{\underline{v}}_2^- = \underline{y}_2 - \underline{x}_2(\underline{x}_2' \underline{x}_2)^{-1} \underline{x}_2' \underline{y}_2 \quad (34)$$

Dann gilt ¹⁾

$$E(\hat{\underline{v}}_2^-) = \underline{x}_3 \gamma_{23} - \underline{x}_2(\underline{x}_2' \underline{x}_2)^{-1} \underline{x}_2' \underline{x}_3 \gamma_{23} + \underline{v}_2 \quad (35)$$

Mit einer Umformulierung nach Schneeweiß ²⁾ kann man für die Gleichung (5.1) auch unter Verwendung der Kurzschreibweise $\delta = (\underline{x}_2' \underline{x}_2)^{-1} \underline{x}_2' \underline{x}_3$ schreiben

$$\underline{y}_1 = \underline{x}_1 \gamma_{11} + (\gamma_{13}/\gamma_{23}) \hat{\underline{v}}_2^- + \underline{x}_2 \gamma_{13} \delta - (\gamma_{13}/\gamma_{23}) \hat{\underline{v}}_2^- + \underline{v}_1 \quad (36)$$

Damit hätte man mit $\sigma^2(\underline{v}_2) = 0$ eine unverzerzte Schätzung für den Koeffizienten γ_{11} der ersten Gleichung (5.1) nach

$$E[(\underline{x}_1, \underline{x}_2, \hat{\underline{v}}_2^-)' (\underline{x}_1, \underline{x}_2, \hat{\underline{v}}_2^-)]^{-1} (\underline{x}_1, \underline{x}_2, \hat{\underline{v}}_2^-)' \underline{y}_1 = (\gamma_{11}, \gamma_{13} \delta, \frac{\gamma_{13}}{\gamma_{23}})' \quad (37)$$

1) Vgl. hierzu Kapitel 2.1.2. S. 18

2) Schneeweiß (1974) S. 106 f.

Neben der nur vorläufig zu Demonstrationszwecken eingeführten Annahme $\sigma^2(\underline{v}_2) = 0$ ist vor allem die Einbeziehung der Variablen $\underline{x}_2$ in den Schätzansatz für eine unverzerzte Schätzung entscheidend: Nur durch $\hat{\underline{v}}_2^-$ und $\underline{x}_2$ wird der Raum aufgespannt, in dem die fehlende Variable $\underline{x}_3$ liegt.

Damit ist dieser Ansatz verschieden von dem von Telser ¹⁾ vorgeschlagenen iterativen SUR - Ansatz, der nach folgendem Schema abläuft :

1. Stufe: Alle Gleichungen des Systems werden mit OLS geschätzt nach

$$\hat{\underline{z}}_i^{(1)} = (\underline{x}'_i \underline{x}_i)^{-1} \underline{x}'_i \underline{z}_i \quad (i=1, \dots, G) \quad (38)$$

Hieraus werden die geschätzten Residuen bestimmt nach

$$\hat{\underline{v}}_i^{(1)} = \underline{z}_i - \underline{x}_i \hat{\underline{z}}_i^{(1)} \quad (39)$$

2. bzw. r-te Stufe: die aus der (r-1)-ten Stufe bestimmten geschätzten Residuen $\hat{\underline{v}}_i^{(r-1)}$ werden in einen OLS - Schätzansatz einbezogen nach

$$(\hat{\underline{z}}, \hat{\underline{a}})_i^{(r)} = [(\underline{x}_i, \hat{\underline{v}}^{(r-1)})' (\underline{x}_i, \hat{\underline{v}}^{(r-1)})]^{-1} (\underline{x}_i, \hat{\underline{v}}^{(r-1)})' \underline{z}_i \quad (40)$$

wobei die Matrix $\hat{\underline{v}}^{(r-1)}$ die geschätzten Residuen der (r-1)-ten Iteration darstellen nach

$$\hat{\underline{v}}^{(r-1)} = (\hat{\underline{v}}_1, \hat{\underline{v}}_2, \dots, \hat{\underline{v}}_{i-1}, \hat{\underline{v}}_{i+1}, \dots, \hat{\underline{v}}_G)^{(r-1)}$$

Wie Telser schreibt, konvergiert dieses Verfahren "beinahe immer" ²⁾ zu festen Schätzwerten $\hat{\underline{z}}_i$, wenn T genügend groß ist. Er zeigt weiterhin, daß die asymptotischen Eigenschaften seines Schätzers mit denen des zweistufigen Zellner-Schätzers übereinstimmen. Wegen vollständig übereinstimmender Ergebnisse ihrer Simulationsrechnungen vermuten Kmenta und Gilbert ³⁾, daß der iterative Telser - Schätzer und der iterative Zellner-

1) Telser (1964)

2) Telser (1964) S. 860 . Eigene Simulationen des Verfassers ergaben ebenfalls diese Aussage.

3) Kmenta & Gilbert (1968) S. 1191

Schätzer " unterschiedliche Wege sind, das selbe Problem der Lösung eines Systems von nichtlinearen Gleichungen zu berechnen" . Diese Übereinstimmung der Ergebnisse konnte auch bei eigenen Simulationen des Verfassers festgestellt werden ¹⁾, sodaß bei der späteren Demonstration der Ergebnisse iterativer SUR - Schätzungen lediglich das Ergebnis des Iterationsverfahrens nach Zellner wiedergegeben wird.

Wie man sehen konnte, war das gleichzeitige Einbeziehen von $\hat{\underline{v}}_2$ und $\underline{x}_2$ die wesentliche Bedingung, unverzerrte Schätzungen für γ_{11} zu erhalten. Der Koeffizient γ_{13} konnte selbstverständlich nicht geschätzt werden, da bei nicht vorhandenen Meßwerten von $\underline{x}_3$ über die Varianz von $\underline{x}_3$ nichts bekannt ist und damit ein der Standardabweichung von $\underline{x}_3$ proportionaler Faktor in γ_{13} willkürlich zu wählen ist.

Rechnerisch einfacher als die zur Abgrenzung gegenüber dem Schätzer nach Telser vorgeführte zweistufige Vorgehensweise ist aber folgender Rechengang: Man setze direkt in der ersten Stufe die Variablen $\underline{y}_2$ und $\underline{x}_2$ als Regressoren ein nach (4) und schätze dann mit OLS die Gleichung (5.1) direkt, sodaß man mit $\sigma^2(\underline{v}_2) = 0$ nun erhält

$$\begin{aligned} E[(\underline{x}_1, \underline{x}_2, \underline{y}_2)' (\underline{x}_1, \underline{x}_2, \underline{y}_2)]^{-1} (\underline{x}_1, \underline{x}_2, \underline{y}_2)' \underline{y}_1 &= (\gamma_{11}, \frac{-\gamma_{22}\gamma_{13}}{\gamma_{23}}, \frac{\gamma_{13}}{\gamma_{23}})' \\ &= (\gamma_{11}, \delta_{11}, \beta_{11})' \end{aligned} \quad (41)$$

Wenn man also eine zweite fehlerfreie Gleichung in dem Modell hätte, deren "Regressoren" mit denen von $\underline{y}_1$ nur $\underline{x}_3$ in der Schnittmenge haben, könnte man wegen der Existenz einer "Definitionsgleichung" für $\underline{x}_3$ den Koeffizienten γ_{11} unverzerrt mit OLS schätzen. Die Annahme einer fehlerfreien

1) Wenn auch das Verhalten während der Iterationen für die beiden Schätzer deutlich verschieden war, so lagen die Unterschiede in den Ergebnisse bei konvergentem Verhalten innerhalb der Rechengenauigkeit des benutzten Rechners.

Definitionsgleichung läßt sich aber für die Schätzpraxis nicht aufrecht erhalten. Sobald man aber einen Fehler in $\underline{y}_2$ zuläßt, liegt wieder ein "Fehler in dem Regressor" - Modell vor, konsistente Schätzungen für γ_{11} sind ohne zusätzliche Informationen nicht zu erlangen. Trotzdem schlägt der Verfasser das folgende Verfahren vor:

Liegt ein Gleichungssystem der Form (22) vor, bei dem folgende über (23) und (24) hinausgehende Annahmen getroffen werden

a) Der Ökonometriker weiß, daß in den o. B. d. A. l' ersten Gleichungen seines Systems die spezifizierten Regressoren X_i ($i=1, \dots, l'$) eine Unterspezifikation dieses Teilmodells in dem Sinne darstellen, daß in allen l' Gleichungen gemeinsam $l = l' - 1$ unbekannte oder nicht meßbare Regressoren $\underline{x}_{\lambda}$ ausgeschlossen wurden mit $\underline{X}^- = (\underline{x}_1^-; \underline{x}_2^-; \dots; \underline{x}_{\lambda}^-; \dots; \underline{x}_1^-)$.

b) Die in den l' ersten Gleichungen spezifizierten Regressoren X_i ($i=2, \dots, l'$) sind für alle l' Gleichungen disjunkt zu X_1 , das heißt die Schnittmenge der spezifizierten Regressoren X_j mit X_1 ist leer.

Wenn dann ein OLS - Regressionsansatz für z. B. die erste Gleichung nach

$$(\hat{\underline{\gamma}}_1, \hat{\underline{\delta}}_1, \hat{\underline{\beta}}_1)' = [(X, Y_1)'(X, Y_1)]^{-1} (X, Y_1)' \underline{y}_1 \quad (42)$$

mit vollem Rang von $[(X, Y_1)'(X, Y_1)]$ und

$$X = (X_1, X_2, \dots, X_{l'}) \quad \text{und} \quad Y_1 = (\underline{y}_2, \underline{y}_3, \dots, \underline{y}_{l'}) \quad (43)$$

und den Annahmen (23) und (24) für $i=1, \dots, l'$ für diese Gleichung ein hohes adjustiertes Bestimmtheitsmaß liefert nach

$$B(\text{adj.}) = 1 - \frac{T-1}{T-K} (1 - R^2) \quad (44)$$

wobei $K' = \sum_{i=1}^{l'} K_i$ die Summe der Regressoren in (42) ist und

R^2 das Bestimmtheitsmaß bzw. das Quadrat des multiplen

Korrelationskoeffizienten darstellt, dann ist dieser Ansatz einem nur mit den X_i ($i=1, \dots, l'$) spezifizierten Ansatz aus drei Gründen vorzuziehen:

1) Aus der Faktoranalyse ist bekannt ¹⁾, daß das Bestimmtheitsmaß R^2 eine untere Schranke für die Kommunalität eines Faktorenmodells darstellt. Das vorliegende Modell aber ist ja nichts anderes als ein Faktorenmodell, bei dem die unbekannten Variablen X_i^- die Faktoren sind, die neben den bekannten X_i ($i=1, \dots, l'$) die Variablen y_i erklären. Daß hier der Sonderfall einer "konfirmatorischen Faktorenanalyse" ²⁾ vorliegt, bei dem die "gemeinsamen" Faktoren X_i^- nicht unkorreliert sein müssen und bei dem $\hat{a}$ priori festgelegt ist, welche "Indikatoren" y_i von welchen "Faktoren" X_i^- abhängig sind, ändert an dieser Aussage nichts: Entscheidend ist die Annahme (23) : $E(\underline{v}_i' \underline{v}_j) = 0$, d. h. die "Einzelrestfaktoren" müssen orthogonal sein. Wenn also eine nach (42) spezifizierte Schätzung ein hohes Bestimmtheitsmaß für alle i ($i=1, \dots, l'$) Gleichungen liefert ³⁾, so ist die Verzerrung der Schätzung wegen des Proximity - Theorems von Wold ⁴⁾ gering.

2) Nach der Beweisführung von McCullum und Wickens ist bei der Verwendung einer Proxyvariablen die Verzerrung des Schätzansatzes immer kleiner als diejenige, die sich bei einem Ansatz ohne die Proxies ergeben würde ⁵⁾.

3) Wenn durch einen nach (42) spezifizierten Ansatz das Bestimmtheitsmaß deutlich zunimmt gegenüber einem nur mit den X_i spezifizierten Ansatz, ist dies für den Ökonometriker

1) Vgl. z. B. Guttman (1940) und vor allem (1957)

2) Vgl. z. B. Weede & Jagodzinsky (1977)

3) Die Berechnung von $B(\text{adj})$ statt B stellt eine vorsichtigeren Aussage sicher, da gilt: $B(\text{adj}) \leq B$ (Vgl. Theil (1971) S.179)

4) Vgl. Kapitel 1 S. 6

5) Vgl. Kapitel 3.4.1. Seite 26

ein deutlicher Hinweis darauf, daß ein großer Anteil der für die $\underline{y}_i$ erklärten Varianz von Variablen stammt, über die er außer dem Wissen ihres Einflusses in den Gleichungen i ($i=1, \dots, l'$) keinerlei Informationen besitzt: Ein deutliches Warnsignal dafür, wie unvollständig die Spezifikation nur mit den X_i ist.

Bei der Suche nach den "fehlenden" Regressoren X_i kann ihm unter Umständen schon der Regressionsansatz (42) helfen. Denn faßt man die "Faktoren" X_i ($i=2, \dots, l'$) und Y_1 mit dem gewonnenen Schätzergebnis $\hat{\underline{\beta}}_1$ und $\hat{\underline{\beta}}_1$ zusammen nach

$$\hat{\underline{D}}_1 = (X_2, X_3, \dots, X_{l'}, Y_1) (\hat{\underline{\beta}}_{11}, \hat{\underline{\beta}}_{12}, \dots, \hat{\underline{\beta}}_{1, l'} - \hat{\underline{\beta}}_1)' \quad (45)$$

so kann diese Dummy- oder Hilfsvariable möglicherweise dem Ökonometriker einen Hinweis darauf geben, welchen (z. B. zeitlichen) Verlauf diese fehlenden Variablen besitzen. und damit möglicherweise die Festlegung eines z. B. wirtschaftlichen Inhalts der Bedeutung dieser Variablen gestatten. Aus diesem Grund werden in den Programmen der im numerischen Anhang vorgeführten Beispiele die Werte von $\hat{\underline{D}}_1$ immer unter der Bezeichnung "Hilfsvariable für $\underline{y}_1$ " ausgedruckt; diese Variablen sind also nicht mit Instrumentalvariablen zu wechseln.

6.3. Konsistenter SURE-Ansatz bei Unterspezifikation

Wie zu Beginn dieses Kapitels gezeigt wurde, können unter den Bedingungen (23) bis (28) Gleichungen eines SUR - Systems dazu dienen, als Instrumentalvariable für Proxy - Variable einen konsistenten OLS - Schätzansatz zu ermöglichen.

Dies ist selbstverständlich auch für den gerade behandelten Fall gemeinsam ausgeschlossener Regressoren möglich, wobei ein solcher Ansatz bei einer großen Anzahl gemeinsam ausge-

schlossener Regressoren nach Meinung des Verfassers allerdings akademischen Charakter annimmt. Für eine nicht zu große Anzahl ausgeschlossener Regressoren bei einem großen Modell kann er aber durchaus seine Berechtigung finden. In Analogie zu (31), (32) und (33) sowie (42) und (43) ist jetzt zu fordern

1) Es liege ein Modell nach (22) vor, das die Bedingungen (23) und (24) erfülle.

2) Der Ökonometriker weiß, daß in den o. B. d. A. ersten $L' \geq 2l + 1$ Gleichungen insgesamt l Regressoren $\underline{x}_{\lambda}$ gemeinsam ausgeschlossen wurden mit $\lambda = 1, \dots, l$.

3) Die in den $l' = l + 1$ ersten Gleichungen spezifizierten Regressoren X_i ($i = 2, \dots, l'$) sind für alle l' Gleichungen disjunkt zu X_1 , die Schnittmenge der spezifizierten Regressoren X_i mit X_1 sei also leer.

Damit ist eine Aufteilung der Gleichungen i ($i = 2, \dots, l'$) in zwei Gruppen möglich nach

$$(X, Y_1)^I = (X_2, X_3, \dots, X_{l'}, Y_2, Y_3, \dots, Y_{l'}) \quad (46)$$

und

$$(X, Y_1)^{II} = (X_{1'+1}, X_{1'+2}, \dots, X_{L'}; Y_{1'+1}, Y_{1'+2}, \dots, Y_{L'}) \quad (47)$$

Da jetzt die Y_i ($i = 2, \dots, l'$) die Proxy-Variablen darstellen, lautet die erste Stufe in Analogie zu (18)

$$\hat{Y}_i = (X, Y_1)^{II} [(X, Y_1)^{II}, (X, Y_1)^{II}]^{-1} (X, Y_1)^{II}, Y_i \quad (48)$$

für $i = 2, \dots, l'$. Anschließend verwendet man diese $\hat{Y}_i$ als Instrumentalvariablen für die Y_i in dem Instrumentalvariablenansatz

$$(\hat{Y}_1, \hat{Y}_2, \hat{Y}_3)' = [(X, \hat{Y}_1)^I, (X, Y_1)^I]^{-1} (X, \hat{Y}_1)^I, Y_1 \quad (49)$$

wobei die Matrix $\hat{Y}_1$ definiert ist nach

$$\hat{Y}_1 = (\hat{Y}_2, \hat{Y}_3, \dots, \hat{Y}_{l'})$$

und $[(X, \hat{Y}_1)^I, (X, Y_1)^I]^{-1}$ sowie $[(X, Y_1)^{II}, (X, Y_1)^{II}]^{-1}$ existieren.

Dieser Ansatz liefert konsistente Schätzungen für $\underline{Y}_1$, da neben den exogenen Variablen X auch die Y_1^{II} wegen (23) die an Instrumentalvariable gestellten Bedingungen erfüllen.

Dieser Ansatz ist ein deutliches Beispiel für die wachsende zusätzliche Information aus dem Gleichungssystem, die das gegenüber dem Proxy-Variablen-Ansatz nach (31) und (32) niedrigere Wissen des Ökonometrikers über die ausgeschlossenen Variablen kompensieren muß: Hatte er dort wenigstens eine (allerdings fehlerbehaftete) Näherung für die fehlenden Variablen, so beschränkt sich sein Wissen im vorliegenden Fall auf die Feststellung, daß in einer bestimmten Teilmenge der Gleichungen seines Modells ein dieser Menge gemeinsamer Satz Regressoren "fehlt".

Eine zusätzliche Ergänzung zu der von beiden Schätzverfahren für den Fall gemeinsam ausgeschlossener Regressoren erhobenen Bedingung der Disjunktivität der l' ersten Regressoren X_i ¹⁾ erweitert unter bestimmten Bedingungen die Anwendbarkeit der Schätzverfahren. Diese Bedingung sorgt ja dafür, daß in keiner der "Regressorgleichungen" ($\underline{Y}_i, X_i$; $i=2, \dots, l'$) der spezifizierte Regressorensatz X_1 auftaucht, also keine "Überlagerung" in dem Ansatz nach (42) oder (49) von X_1 und den X_i ($i=2, \dots, l'$) stattfinden kann. Nur dadurch ist bei dem Ansatz (49) die konsistente OLS - Schätzung aller Elemente des Vektors $\underline{Y}_1$ möglich. Wie später auch an einem kleinen Beispiel gezeigt wird, kann man durch "stückweises" Schätzen des Vektors $\underline{Y}_1$ diese Bedingung einschränken; der Einfachheit halber sei dies nur für den Ansatz (49) gezeigt. Es gilt nämlich: Jedes Element Y_{1j} des Vektors $\underline{Y}_1$ kann mit dem Ansatz (49)

1) X_1 sei also disjunkt zu $X_2, \dots, X_{l'}$; vgl. S. 70 bzw. 73

konsistent geschätzt werden, wenn der zugehörige Vektor $\underline{x}_{1j}$ in keiner der Matrizen X_i ($i=2, \dots, l'$) enthalten ist, wenn also die in X_i ($i=2, \dots, l'$) spezifizierten Regressoren disjunkt zu $\underline{x}_{1j}$ sind. Dies ist aus (49) sofort zu sehen, denn wenn die (fehlerbehafteten) "Definitionsgleichungen" für X^- in (46) den Raum der ausgeschlossenen Variablen X^- ohne $\underline{x}_{1j}$ aufspannen können, kann keine Überlagerung in (49) stattfinden. Somit kann der Vektor $\underline{\gamma}_1$ dann "stückweise" konsistent mit dem Ansatz (49) geschätzt werden, wenn gilt:

Existiert zu jedem Vektor $\underline{x}_{1j}$ ($j=1, \dots, K_1$) ein Satz von "Definitionsgleichungen" der ausgeschlossenen Regressoren X^- nach (46), der l Elemente $\underline{y}_i$ ($i \neq 1$) und X_i ($i \neq 1$) enthält, die zu $\underline{x}_{1j}$ disjunkt sind, und bezeichnet man diesen Regressorensatz mit $(X, Y_1)^{Ij}$ und den zur ersten Gleichung disjunkten restlichen Satz an Definitionsgleichungen aus L' mit $(X, Y_1)^{IIj}$, so kann man mit maximal j sukzessiven Schätzungen nach folgendem Schätzansatz den gesamten Vektor $\underline{\gamma}_1$ konsistent schätzen

1. Stufe zur Schätzung des j -ten Elementes aus $\underline{\gamma}_1$ nach

$$\underline{\hat{\gamma}}_1^j = (X, Y_1)^{IIj} [(X, Y_1)^{IIj}, (X, Y_1)^{Ij}]^{-1} (X, Y_1)^{Ij}, \underline{y}_1^j \quad (50)$$

mit $\underline{y}_1^j$ aus Y_1^j . Diese $\underline{\hat{\gamma}}_1^j$ werden in der zweiten Stufe als Instrumentalvariablen für die $\underline{y}_1^j$ verwendet in dem Instrumentalvariablenansatz nach

$$(\underline{\hat{\gamma}}_1, \underline{\hat{\delta}}_1, \underline{\hat{\beta}}_1)^{(j)} = [(X, \underline{\hat{\gamma}}_1)^{Ij}, (X, Y_1)^{Ij}]^{-1} (X, \underline{\hat{\gamma}}_1)^{Ij}, \underline{y}_1 \quad (51)$$

Dann ist der Vektor $\underline{\hat{\gamma}}_1 = (\hat{\gamma}_{11}^{(1)}, \hat{\gamma}_{12}^{(2)}, \dots, \hat{\gamma}_{1K_1}^{(K_1)})$ eine konsistente Schätzung für $\underline{\gamma}_1$ ①.

Somit existieren nun neben den bekannten Schätzverfahren für SUR - Systeme zwei bzw. drei neue Schätzansätze, die aber von

① Die wohl notgedrungen überladene Nomenklatur wird sicher durch die nachfolgenden Beispiele veranschaulicht.

den "klassischen" SUR - Schätzern durch eine völlig unterschiedliche Zielsetzung bzw. zugrundeliegende Modellvorstellung unterschieden werden können. Dies soll nun an einer Reihe von kleinen numerischen Beispielen veranschaulicht werden, von denen die ersten konstruiert und damit sowohl überschaubar wie auch überprüfbar sind, während das letzte Beispiel auf einen Anwendungsfall aus der Literatur zurückgreift.

6.4. Beispiele

In den folgenden Beispielen wird jeweils nur auf besonders markante Ergebnisse oder Eigenheiten der im numerischen Anhang vollständig aufgeführten Rechnungen Bezug genommen, um den Text nicht zu überfrachten. Wenn in diesen Beispielen neben den neu eingeführten Schätzverfahren auch außer den OLS - Schätzungen die SURE - Ergebnisse nach dem zweistufigen und dem iterativen Zellner-Schätzer aufgeführt werden, so muß hier betont werden, daß diese Schätzverfahren ja nicht für die vorliegenden Fälle konstruiert wurden und deshalb im allgemeinen schlechtere Ergebnisse liefern müssen. Im Vorgriff auf die Einzelbeispiele und als Ergebnis von rund 200 nicht aufgeführten Simulationen, in denen die laut Rao ¹⁾ für die Verzerrung des Zellner-Schätzers entscheidenden Größen systematisch variiert wurden, kann hier schon gesagt werden, daß die Verzerrung des iterativen Zellner - Schätzers in keinem Fall größer war als diejenige des OLS-Schätzers, jedoch viele Simulationen verblüffend gute Ergebnisse lieferten. Dies kann ohne weiteres als Widerspruch zu der Auffassung von Maddala gewertet werden, im Fall gemeinsam ausgeschlossener Regressoren "versage" der Zellner-Schätzer.²⁾

1) Rao (1974) S. 108 ; dort steht auch der mißverständliche Satz: "...the bias of the SURE estimator is generally larger than the bias of the OLS estimator."

2) Vgl. S. 56

6.4.1. Beispiel des Anhangs Seite A.V.

Dieses Modell stellt einen Anwendungsfall für einen Schätzansatz nach (12) dar, wobei als Besonderheit $\underline{x}_1 = \underline{x}_2$ gilt. Deshalb kann hier ein SUR-Ansatz nach Zellner keine besseren Schätzungen als OLS liefern ¹⁾. Bemerkenswert an den ansonsten für sich selbst sprechenden Ergebnissen ist die Korrelation zwischen den Residuen $\underline{v}_1$ und $\underline{v}_2$ ²⁾, die im korrekt spezifizierten Modell Null ist, während sie im fehlspezifizierten Proxy - Ansatz mit .9638 sehr hoch ist. Dies ist ein Beispiel dafür, daß eine durch gemeinsame Fehlspezifikation mehrerer Gleichungen nichtverschwindende Kovarianz zwischen den geschätzten Gleichungsresiduen nicht nur durch einen gemeinsamen Ausschluß von Regressoren aus der Spezifikation entstehen kann, sondern auch durch die gemeinsame Einbeziehung des falschen Regressortyps, also einer Proxy. Daß die an den "t" - Werten vergleichbare Effizienz des EIV - Ansatzes gegenüber der Originalspezifikation nur geringfügig geringer ist, liegt z. B. für die Gleichung (V.7) an dem für dieses Modell geringen Varianzanteil von $\sigma^2(\underline{v}_2)$ an $\sigma^2(\underline{y}_2)$: Damit wird die Korrelation $r(\underline{y}_2, \underline{x}_2)$ nur wenig kleiner als $r(\underline{y}_2, \underline{x}_2)$ und $\underline{y}_2$ ist damit eine "gute" Instrumentalvariable für $\underline{x}_2$ ³⁾.

6.4.2. Beispiel des Anhangs Seite A.VI.

Auch hier liegt ein Anwendungsfall für den Schätzansatz nach (12) vor, jetzt allerdings mit $\underline{x}_1 \neq \underline{x}_2$. Die Effizienzsteigerung von z. B. Gleichung (VI.7) über (VI.8) zu (VI.9) wurde schon auf Seite 63 kommentiert. Im Gegensatz zum Beispiel A.V. liefert jetzt aber der Zellner-Schätzer von OLS abweichende

1) Vgl. z. B. Johnston (1972) S. 240

2) Die Residuen $\underline{v}_i$ werden in den Programmausdrucken aus programmiertechnischen Gründen mit u bezeichnet.

3) Vgl. Dhrymes (1970) S. 298

Ergebnisse. Schon die Resultate des zweistufigen Zellner - Schätzers sind überraschend gut im Sinne kleiner Verzerrung (Gleichungen (VI.13) und (VI.14)), die Ergebnisse des iterativen Zellner - Schätzers (Gleichungen (VI.15) und (VI.16)) aber sind geradezu hervorragend. Ein Effizienzvergleich mit dem EIV - Ansatz der Gleichungen (VI.7) bis (VI.12) ist wegen der Verzerrung des Zellner - Schätzers nicht sinnvoll.

Bei den Ergebnisausdrücken der Zellner - Schätzungen ist zu beachten, daß bei der Berechnung von s^2 der Nenner " F_{ij} "¹⁾ in Anpassung an Theil²⁾ und Zellner³⁾ zu $F_{ij} = T$ gewählt wurde. Aus demselben Grund werden auch die Standardabweichungen der Regressionskoeffizienten mit der Bezeichnung "Sigma" ausgegeben und keine "t" - Werte. Bei den Ausdrücken des EIV - Ansatzes nach (12) ist zu beachten, daß die Angabe von "t" - Werten keine t-Verteilung vortäuschen soll: Da für den EIV - Ansatz lediglich die Konsistenz gezeigt wurde, liegt hier keine t-Verteilung, sondern mit $T \rightarrow \infty$ eine asymptotische Normalverteilung⁴⁾ vor. Zur Vergleichbarkeit mit den OLS-Ergebnissen und nach dem Vorbild von Dhrymes⁵⁾ werden aber auch beim Vorliegen eines lediglich konsistenten Schätzers "t" - Werte berechnet und ausgedrückt, wobei die Anführungszeichen den Vorbehalt lediglich asymptotischer Gültigkeit ausdrücken sollen.

1) Vgl. Seite 53 f.

2) Theil (1971) S. 295 und S. 321

3) Zellner (1963) S. 980

4) Dhrymes (1970) S. 198 ff.

5) Dhrymes (1970) S. 533. Dort gibt er für eine Zweistufige Least - Squares - Schätzung ebenfalls "t" - Werte an, obwohl er (auf S. 198 ff.) zeigt, daß die Koeffizienten dieser Schätzung lediglich asymptotisch normalverteilt, aber nicht t-verteilt sind. Der Grund liegt wohl in dem Wunsch, die Ergebnisse mit OLS - Ergebnisse vergleichen zu können, für die immer t-Werte angegeben werden.

6.4.3. Beispiel des Anhangs Seite A.VII.

Hier liegt ein Anwendungsfall für den (inkonsistenten) Schätzansatz nach (42) vor : Aus einem Modell nach (22), das die Bedingungen (23) erfüllt, ist in allen Gleichungen die Variable $\bar{x}_3$ ausgeschlossen. Das Ausgangsmodell ist in den Gleichungen (VII.1) und (VII.2) wiedergegeben, die nach Ausschluß von $\bar{x}_3$ resultierenden OLS-Schätzungen unter (VII.3) und (VII.4) sowie die Zellner - Schätzungen unter (VII.5) bis (VII.8). Die Gleichungen (VII.9) und (VII.10) geben die Schätzungen nach dem Schätzansatz (42) wieder, wobei $\hat{D}_i$ aus (45) im Programmausdruck mit "ZZ" bezeichnet wird und als "Hilfsvariable" auch ausgedruckt wird. Zur Veranschaulichung der in "ZZ" subsumierten Größen, also für (VII.9) z. B. "ZZ" = $\hat{\sigma}_1 \bar{x}_2 + \hat{\beta}_1 \bar{y}_2$, sind diese in den Gleichungen (VII.11) und (VII.12) disaggregiert wiedergegeben.

Vergleicht man alle vier Schätzansätze, so kann man bezüglich der Verzerrung eine Rangfolge absinkender Verzerrung aufstellen: Die erhebliche Verzerrung von OLS wird von dem zweistufigen Zellner-Schätzer um ein Geringes, von dem iterativen Zellner-Schätzer aber schon sehr deutlich unterboten.¹⁾ Die geringste Verzerrung liefert der Ansatz nach (42), wobei deutlich sichtbar wird, daß wegen $\sigma^2(\bar{y}_1) > \sigma^2(\bar{y}_2)$ die Schätzung (VII.10) eine größere Verzerrung als (VII.9) aufweist: $\bar{y}_1$ ist in Verbindung mit $\bar{x}_1$ eine schlechtere Proxy für $\bar{x}_3$ als $\bar{y}_2$ in Verbindung mit $\bar{x}_2$. Die "t"-Werte haben wegen der Verzerrung der Schätzungen selbstverständlich keine exakte Aussagefähigkeit, auch sie gelten höchstens approximativ als Maß für die Signifikanz der Koeffizienten. Aber trotz

1) Daß der iterative Zellner-Schätzer allerdings 34 Iterationen bis zur Konvergenz benötigte, kann seine Anwendung bei deutlich größeren Modellen aus Rechenzeitgründen ausschließen.

der deutlichen Verminderung der Verzerrung sei an dieser Stelle noch einmal betont, daß an erster Stelle die von einer nicht von dem Datensatz inspirierten (z. B. volkswirtschaftlichen) Hypothese getragene Auffassung des Ökonometrikers stehen muß, daß aus einer von ihm genau formulierten Teilmenge der Systemgleichungen ein Satz von Regressoren "fehlt". Der Schätzansatz (42) kann bei einem nicht durch eine Hypothese abgestützten "Zusammenwerfen" von Gleichungen nur zufällig, wenn überhaupt, sinnvolle Ergebnisse liefern; als Instrument zur empirischen Falsifikation von Hypothesen aber ist er dann nicht geeignet. Und dies ist ja letztlich die Aufgabe eines jeden ökonometrischen Modells. Vor allem soll auch vor dem falschen Umkehrschluß gewarnt werden, daß eine Gleichung des Systems die auf Seite 70 formulierten Bedingungen (vor allem die Bedingung a) !) erfüllt, wenn sie nach Durchprobieren vieler anderer Alternativen endlich bei ihrer Verwendung als "Definitionsgleichung" in (42) dafür sorgt, daß bestimmte Koeffizienten in $\hat{\gamma}_1$ das "richtige", d. h. hypothesenkonforme Vorzeichen erhalten.

6.4.4. Beispiel des Anhangs Seite A.VIII.

Dieses Modell entspricht in seinem Aufbau vollständig dem Modell A.VII. , benutzt allerdings einen veränderten Datensatz. Es steht stellvertretend für eine große Anzahl aus den rund 200 Simulationen, bei denen der iterative Zellner-Schätzer geradezu verblüffend gute Ergebnisse lieferte, obwohl die von OLS gelieferten "Startwerte" stark verzerrt waren ¹⁾. Die starke Übereinstimmung der iterativen Zellner-Schätzungen mit den Schätzergebnissen nach (42) machen jetzt auch deutlich, warum auf eine Abgrenzung gegenüber dem

1) Vgl. z. B. die Gleichungen VIII.6 und VIII.10 !

iterativen SUR - Schätzer nach Telser ¹⁾ so großer Wert gelegt wurde, da ja bei iterativen Schätzverfahren des vorgestellten Typus' bei einem fehlspezifizierten Modell nicht bekannt ist, ob und vor allem zu welchem Wert sie konvergieren. Daß dabei die langsame Konvergenz bei großen Modellen unter Umständen die Rechenzeit unvertretbar anwachsen läßt, wurde schon bei dem Beispiel VII. angedeutet.

6.4.5. Beispiel des Anhangs Seite A.IX.

Dieses Modell ist ein Anwendungsbeispiel für den Schätzansatz (49): Die Variable $\underline{x}_3$ ist in allen drei Gleichungen des Modells enthalten, sei aber für einen Schätzansatz nicht verfügbar. Neben der Beschreibung des "wahren" Modells in den Gleichungen (IX.1) bis (IX.3) sind mit den Gleichungen (IX.4) bis (IX.15) die Ergebnisse der inkonsistenten Schätzungen mit OLS, den beiden Zellner-Schätzern und dem Schätzansatz nach (42) wiedergegeben. Auch hier wieder sind die erstaunlich guten Ergebnisse des iterativen Zellner-Schätzers hervorzuheben, die allerdings theoretisch nicht zu untermauern sind. Die konsistenten Schätzungen nach (49) sind schließlich in den Gleichungen (IX.16) bis (IX.19) wiedergegeben. Betrachtet man z. B. die Gleichung (IX.16), so wird das Schätzverfahren deutlich: Die synthetische Variable $(\underline{y}_2, \underline{x}_2)$ dient als mit dem Fehler $\underline{v}_2$ behaftete Proxy-Variable für die fehlende Variable $\underline{x}_3$ nach dem Beispiel der Gleichung (4). Da die Spezifikation der zweiten Modellgleichung (IX.2) kein $\underline{x}_2$ enthält, ist die Disjunkтивität zur Spezifikation von (IX.1) gesichert, eine "Überlagerung" von $\underline{x}_1$ -Anteilen in dem Schätzansatz (49) kann also nicht stattfinden. Da eine solche Restriktion für die als Instrumentalvariable

1) Vgl. Seite 68

dienende Gleichung (IX.3) nicht gilt ¹⁾, berührt das Auftreten von $\underline{x}_1$ in (IX.3) bzw. (IX.16) nicht die konsistente Schätzung von γ_{11} in (IX.16). Analog ist damit auch in Gleichung (IX.17) die Variable $(\underline{y}_1, \underline{x}_1)$ als Proxy-Variable für $\underline{x}_3$ geeignet, während wiederum $\underline{y}_3$ als Instrumentalvariable für diese Proxy die an eine Instrumentalvariable gestellten Bedingungen erfüllt.

Die gegenüber dem ursprünglichen Modell auftretende Abweichung der Schätzungen von s^2 ist keine "Fehlschätzung", sondern ist auf die durch die Proxyvariablen zusätzlich eingeführten Regressoren $\underline{x}_2$ (in (IX.16)) bzw. $\underline{x}_1$ (in (IX.17)) zurückzuführen. Dadurch ist bei dem zur Vergleichbarkeit mit OLS verwendeten Schätzer für s^2 nach (17) mit $Fr = T - K_i$ der Wert K_i um die von den "Proxy-Gleichungen" eingebrachte Anzahl von exogenen Variablen zu groß ²⁾; eine Umrechnung für z. B. die Gleichung (IX.16) nach $s_1^2 = s_1^2 \cdot \frac{T}{T - K_i} = s_i^2 \cdot \frac{T}{T - K_i}$ mit $s_i^2 =$ ausgedruckter s^2 -Wert, K_i = Anzahl der Regressoren des korrekt spezifizierten Modells und K'_i = Anzahl der Regressoren des durch die Proxygleichungen spezifizierten Modells ergibt aber die vergleichbaren Schätzwerte für s_1^2 ³⁾.

Die Gleichung (IX.18) schließlich ist ein Anwendungsbeispiel für eine "gestückelte" Schätzung nach dem Ansatz (51) : Da in der korrekt spezifizierten Gleichung (IX.3) sowohl $\underline{x}_1$ wie auch $\underline{x}_2$ enthalten sind, führt die Wahl von z. B. der ersten Modellgleichung als Proxy für $\underline{x}_3$ zu einer Verzerrung von $\hat{\gamma}_{31}$, da ja die erste Gleichung $\underline{x}_1$ benötigt, um den (hier eindimensionalen) Raum der fehlenden Variablen aufzuspannen. Damit

1) Vgl. z. B. Kap. 3.5.1 oder das Vorgehen bei Two Stage Least Squares.

2) Dieses Problem entfällt natürlich bei der Wahl $Fr = T$.

3) Diese Umrechnung muß immer von dem Ökonometriker durchgeführt werden, da ein Programm nicht zwischen "echten" und "Hilfs-" Regressoren unterscheiden kann.

ergibt sich hier $\hat{\gamma}_{31} = \gamma_{31} - \frac{\gamma_{33}\gamma_{11}}{\gamma_{13}} = \gamma_{31} - \hat{\beta}_{31}\gamma_{11} = \frac{2}{3}$.

Der Koeffizient $\hat{\gamma}_{32}$ aber ist eine konsistente Schätzung für γ_{32} . Entsprechend ist in der Gleichung (IX.19) zwar der Koeffizient $\hat{\gamma}_{32}$ eine inkonsistente Schätzung für γ_{32} , nicht jedoch der Schätzer $\hat{\gamma}_{31}$, der eine konsistente Schätzung für γ_{31} darstellt. Damit kann man eine konsistente Schätzung für γ_3 aus $\hat{\gamma}_{31}$ aus (IX.19) und $\hat{\gamma}_{32}$ aus (IX.18) zusammensetzen; der Koeffizient der fehlenden Variablen kann natürlich nicht geschätzt werden.

Die Gleichungen (IX.20) bis (IX.22) sollen beispielhaft die unterschiedlichen Forderungen veranschaulichen, die an die zur Schätzung benutzten "Hilfsgleichungen" je nach ihrer Verwendung gestellt werden: In Gleichung (IX.20) wird verdeutlicht, daß $\underline{y}_3$ auch ohne $\underline{x}_1$ und $\underline{x}_2$ als Instrumentalvariable für $\underline{y}_2$ verwendet werden kann, solange $\underline{y}_3$ und $\underline{y}_2$ eine nichtverschwindende Kovarianz haben und $\underline{y}_2$ und $\underline{y}_3$ unkorreliert sind. Dies aber ist durch das Ausgangsmodell gewährleistet. Gleichung (IX.21) und (IX.22) zeigen hingegen, daß $\underline{y}_2$ nur als Proxy für $\underline{x}_3$ dienen kann, wenn gleichzeitig $\underline{x}_2$ (wegen (IX.2)!) zum Aufspannen des zur Abbildung von $\underline{x}_3$ dienenden Raumes in den Schätzansatz einbezogen werden. Unterbleibt dies, so sind die Schätzungen für γ_{11} inkonsistent; dies kann auch nicht durch Verwendung von $\underline{y}_3$ mit $\underline{x}_1$ und $\underline{x}_2$ als Instrumentalvariablen für $\underline{y}_2$ geheilt werden, sondern hat im vorliegenden Beispiel sogar eine Vergrößerung der Verzerrung zur Folge.

6.4.6. Beispiel des Anhangs Seite A.X.

Dieses Modell ist nicht konstruiert, sondern soll an einem realen ¹⁾ Modell aufzeigen, wie aus unterschiedlichen Modellvorstellungen des Ökonometrikers und den daraus folgenden ²⁾ Schätzansätzen deutlich divergierende Resultate gefunden werden können bei einem vorgegebenen Datensatz.

Das Modell, das von Zellner ³⁾, Theil ⁴⁾ sowie von Kmenta und Gilbert ⁵⁾ zur Demonstration der SUR-Ansätze von Zellner verwendet wird, geht auf eine Veröffentlichung von Boot und de Witt ⁶⁾ zurück, die zur Schätzung von Investitionsungleichungen folgenden Ansatz durchführen:

Nach einer nicht veröffentlichten These von Grunfeld ⁷⁾ sind nicht die (beobachteten) Gewinne entscheidend für die Investitionen einer Firma, sondern die Gewinnerwartungen. Da diese eine nicht meßbare Variable darstellen, schlägt er als Substitute hierfür den "Marktwert einer Firma" vor, d. h. den Börsenwert der in Umlauf befindlichen Anteilscheine ⁸⁾.

Diese Substitution ist wohl ⁹⁾ auf die Erfahrung begründet, daß Börsenkurse die escomptierten Gewinne aus der Sicht der Börsenteilnehmer darstellen und im wesentlichen mit den Gewinnerwartungen der die Investitionen festlegenden Entscheidungsträger in den Firmen übereinstimmen. Weiter führt Grunfeld den "gewünschten Kapitalstock" zum Zeitpunkt t ein, für den er eine linear wachsende Funktion von dem Marktwert annimmt

1) Dies muß nicht unbedingt realistisch bedeuten.

2) Diese Reihenfolge ist zwingend!

3) Zellner (1962)

4) Theil (1971) S. 295 ff.

5) Kmenta & Gilbert (1968) S. 1199

6) Boot & de Witt (1960)

7) Grunfeld (1958)

8) Da amerikanische Aktiengesellschaften im Gegensatz zu deutschen AG's eigene Aktien besitzen dürfen, ist der Zusatz "im Umlauf befindlich" notwendig.

9) Da die Dissertationsarbeit von Grunfeld nicht veröffentlicht ist, ist der Verfasser auf eigene Vermutungen angewiesen.

nach der Beziehung

$$C_t^+ = c_1 + c_2 F_t \quad (52)$$

mit C_t^+ = gewünschter Kapitalstock zum Zeitpunkt t

und F_t = Marktwert zum Zeitpunkt t . Wenn man mit

C_t den zum Zeitpunkt t vorhandenen Kapitalstock bezeichnet,

dann ist $c_1 + c_2 F_t - C_t$ die gewünschte Nettoinvestition.

Mit der Annahme, daß lediglich ein konstanter Bruchteil q_1

der gewünschten Nettoinvestition im Verlauf des Jahres

zwischen den Zeitpunkten t und $t + 1$ realisiert wird, er-

hält man für die realisierte Nettoinvestition dieses Zeit-
raums

$$q_1(C_t^+ - C_t) = q_1 c_1 + q_1 c_2 F_t - C_t \quad (53)$$

Da aber durch die schließlich angestrebte Investitions-

gleichung die Bruttoinvestition einschließlich Ersatz-

investitionen, Erhaltungsinvestitionen und Reparaturen

beschrieben werden soll und mit der Annahme, daß diese Größen

einen konstanten Anteil q_2 des vorhandenen Kapitalstocks

ausmachen, erhält man schließlich

$$I_{t+1} = q_1 c_1 + q_1 c_2 F_t + (q_2 - q_1) C_t \quad (54)$$

wobei I_{t+1} die Bruttoinvestition in dem auf den Zeitpunkt

t folgenden Jahr darstellt.

Dies stellt die Spezifikation für einen Schätzansatz dar,

wobei leider über die zu erwartenden Koeffizienten kaum

etwas gesagt wird ¹⁾. Zu den Größen q_1 , q_2 und c_2 sagen

Boot und de Witt lediglich, daß sie plausiblerweise zwischen

Null und Eins liegen müssen; zu q_2 äußert Grunfeld die Ver-

mutung, daß q_2 unter 0.3 liegen wird. Im Vorgriff auf ihre

1) Hier wird also die Regression nicht als Hypothesentest verwendet, sondern als Approximationsverfahren im Sinne eines "best fitting" - Verfahrens.

Ergebnisse der OLS - Schätzungen merken Boot und de Witt an, daß der Koeffizient ($q_2 - q_1$) positiv sein kann und ergänzen fast entschuldigend: "This may seem strange for an investment equation, but it must be remembered that the dependent variable is not net investment but gross investment including maintenance and repairs ". Hierzu muß angemerkt werden, daß der untersuchte Zeitraum die Jahre 1935 bis 1954 umfaßt; für diesen Zeitraum aber ist zumindest fraglich, ob in den Vorkriegsjahren und Kriegsjahren für die untersuchten Firmen das gleiche Investitionsverhalten wie in den Boomjahren nach dem zweiten Weltkrieg galt. Vor allem für die Nachkriegsjahre sind Zweifel angebracht, ob bei der enormen Ausweitung des Welthandels und der Erschließung neuer Märkte der Anteil der Erhaltungsinvestitionen am Kapitalstock tatsächlich den Anteil der gewünschten Nettoinvestition am Kapitalstock überwog. Von insgesamt 10 durch Boot und de Witt erfaßten amerikanischen Firmen wurden von Zellner, Theil sowie von Kmenta und Gilbert die Firmen General Electric und Westinghouse für ihre Demonstration des Zellner-Schätzers ausgewählt; als Grund hierfür gibt Theil ¹⁾ an, daß beide Firmen in dem gleichen Industriezweig angesiedelt sind und deshalb zu vermuten ist, "that there are common ²⁾ factors that affect the investment decisions of both firms quite apart from a common development of the two explanatory variables, F_{-1} , F'_{-1} , and C_{-1} , C'_{-1} ³⁾. Since these factors are neglected in the equation, they have to be analyzed in terms of disturbances ." Genau dies

1) Theil (1971) S. 296

2) Die Hervorhebungen wurden von Theil vorgenommen

3) Die vom Verfasser auch in seinen Programmausdrucken übernommene Nomenklatur von Theil bezeichnet die zu General Electric gehörenden Variablen mit F_{-1} und C_{-1} , die zu Westinghouse gehörenden mit F'_{-1} und C'_{-1} .

aber ist die in diesem Kapitel vom Verfasser vorgetragene Vorstellung, die allerdings dann ja zu vom Zellner-Schätzer abweichenden Schätzansätzen führte.

Die Ergebnisse der verschiedenen Schätzungen sind in den Gleichungen (X.1) und (X.2) für OLS, (X.3) und (X.4) für den zweistufigen Zellner-Schätzer und in (X.5) und (X.6) für den iterativen Zellner-Schätzer wiedergegeben und stimmen natürlich mit den Ergebnissen der Schätzungen von Zellner, Theil, Kmenta und Gilbert bzw. Boot und de Witt überein ¹⁾. Entsprechend den von Theil vorgetragenen Überlegungen zeigen die Residuen aller drei Schätzansätze eine hohe Korrelation, was zur Veranschaulichung auf Seite A.X.3 für die Residuen von OLS und dem iterativen Zellner-Schätzer in kleinen Bildern ²⁾ wiedergegeben wird. Bemerkenswert an den Ergebnissen der drei Schätzansätze ist erstens die doch hohe Übereinstimmung der Ergebnisse; die Interdependenz der Gleichungen im Sinne Theils führt lediglich zu geringen Veränderungen der Schätzergebnisse nach Zellner gegenüber OLS. Zweitens ist die deutliche positive Autokorrelation in den Gleichungen (X.1) und (X.2) auffallend, wofür nach Schneeweiß ³⁾ eine Fehlspezifikation beider Gleichungen verantwortlich gemacht werden kann ⁴⁾. Dies deutet nach Meinung des Verfassers alles auf eine Unterspezifikation

- 1) Soweit sie die jeweiligen Schätzer überhaupt anwendeten. Abweichungen bei den Sigma-Werten sind durch die unterschiedliche Wahl von "Fr" bei der Schätzung von s^2 zurückzuführen.
- 2) Da der verwendete Rechner keinen Plotteranschluß besitzt, mußten diese Darstellungen über den Drucker mit seinem zwangsläufig niedrigeren Auflösungsvermögen erfolgen.
- 3) Schneeweiß (1974) S. 188 f.
- 4) Da bei einer Fehlspezifikation nach Fehlertyp II (Vgl. S.18) die Größe $\frac{\hat{v}}{\hat{v}} \cdot \frac{\hat{v}}{\sigma_v^2}$ keiner χ^2 -Verteilung, sondern einer nichtzentralen χ^2 -Verteilung mit Nichtzentralitätsparameter $\lambda = \gamma_x^2(x_2', (x_1'x_1)x_2)/\sigma_v^2$ folgt (Vgl. z. B. Ramsey (1969) S.353), kann auch der Durbin-Watson Test nicht mehr exakt gelten.

beider Gleichungen dieses Modells hin, worauf die neu eingeführten Schätzansätze "zugeschnitten" sind.

Für einen Ansatz nach (19) fehlt allerdings die Grundlage, da dieser das Auftreten der selben Proxy-Variable in mehreren Gleichungen verlangt, die von Grunfeld eingeführten Proxyvariablen F bzw. F' aber für die beiden Gleichungen unterschiedliche Variablen sind.

Nach den Überlegungen von Theil bietet sich aber der zwar inkonsistente ¹⁾, aber gezielt auf das vorliegende Schätzproblem ausgelegte Schätzansatz (42) an, wenn man entweder annimmt, daß lediglich eine gemeinsame Variable als Regressor ausgeschlossen wurde oder daß der "common factor" mehrerer ausgeschlossener Variablen diese zu einem großen (Varianz-) Anteil erklärt. Die Ergebnisse dieses Ansatzes nach (42) sind in den Gleichungen (X.7) und (X.8) wiedergeben, wobei in der Hilfsvariablen "ZZ" der beiden Gleichungen jeweils alle Variablen, also Endogene wie Exogene, der komplementären Gleichung zusammengefaßt sind. Aus statistischer Sicht sprechen die Schätzergebnisse für den Ansatz, denn sowohl das starke Ansteigen des adjustierten Bestimmtheitsmaßes wie auch die Veränderung des Durbin-Watson-Koeffizienten in den Ablehnungsbereich der Autokorrelations-Nullhypothese ²⁾ lassen vermuten, daß die in dem Ansatz nach (42) durchgeführte Verwendung einer beiden Gleichungen fehlenden, gemeinsamen Variablen für eine (fast) korrekte Abbildung des Modells nötig war. Um welche Variable es sich hierbei handeln könnte, ist für den Ökonometriker u. U. aus dem Verlauf der Hilfsvariable "ZZ" zu erkennen, weshalb

- 1) Es darf daran erinnert werden, daß bei Zutreffen der Theil'schen Überlegungen der Zellner-Schätzer ebenfalls inkonsistent ist, worauf leider keiner der Verwender dieses Demonstrationsmodells hinweist.
- 2) Wegen der Inkonsistenz des Schätzers (42) gilt der Durbin-Watson-Test natürlich nicht exakt, kann aber eine "Tendenz" angeben.

neben dem Abdruck dieser Werte auf Seite A.X.4 auch der graphische Verlauf der Werte in kleinen Bildern wiedergegeben wurde. Der zur Abbildung von I bzw. I' von "ZZ" gelieferte Beitrag ist neben der Angabe der "t"-Werte ¹⁾ in den einzelnen Gleichungen auch in den Abbildungen auf Seite A.X.5 verdeutlicht, wo über den OLS-Residuen als Abszisse die Werte der Hilfsvariablen "ZZ" als Ordinate für die beiden Gleichungen aufgetragen wurde.

Aus der interpretatorischen Sicht z. B. eines hypothesentestenden Volkswirtes bereiten die Vorzeichen der Schätzergebnisse für die bei C_{-1} bzw. C'_{-1} stehenden Koeffizienten Schwierigkeiten: Beide Koeffizienten "kippen" von bei allen anderen Schätzansätzen erhaltenen positiven Vorzeichen auf negative Werte um, wobei allerdings die "t"-Werte nur dem bei C'_{-1} stehenden Koeffizienten eine Signifikanz gegen Null zusprechen ¹⁾. Da ohne Einsicht in die Theorie von Grunfeld keine Aussagen darüber gemacht werden können, ob diese Theorie Erwartungen über diese Vorzeichen beinhaltet und welche, kann hier nur der dem Wesen der Ökonometrie widersprechende Weg versucht werden, die Schätzergebnisse plausibel zu machen. Dies darf aber nicht als ein nachträgliches "Nachschieben" einer Theorie verstanden werden in dem Sinne, daß "die Daten die Wahrheit sprechen" und sich eine Theorie daran auszurichten habe; Daten können nicht besser sein als das zugrundegelegte Modell und der modelladäquate Schätzansatz. Mit dieser Einschränkung können für negative Koeffizienten bei C_{-1} bzw. C'_{-1} folgende Überlegungen herangezogen werden: Die Investitionspolitik derart großer Firmen wie der beiden

1) Auch hier sind die "t"-Werte lediglich als Indikatoren, nicht aber als exakte Testgrößen zu verstehen, da ja ein inkonsistenter Schätzer vorliegt.

untersuchten wird sicher sehr stark von dem allgemeinen Konjunkturverlauf geprägt sein, eine azyklische oder gar antizyklische Investitionspolitik ist kaum zu erwarten. Diese Übereinstimmung im Investitionsverhalten stellt die Abbildung (X.G) auf Seite A.X.5 dar. Ein schlechter Konjunkturverlauf aber spiegelt sich meist in einem schlechten Auslastungsgrad der Anlagen nieder, der um so niedriger ist, je höher der - in früheren Boomphasen aufgebaute - Kapitalstock einer Firma ist. Steigen nun die globalen, also nicht firmenspezifischen Gewinnerwartungen in einer Volkswirtschaft, so wird vor einer Erweiterungsinvestition zuerst einmal der vorhandene Anlagenstock bis zu einer maximalen Auslastung ausgenutzt werden. Dies gilt vor allem dann, wenn die Konjunkturphasen nur kurze Dauer besitzen und eine Erweiterungsinvestition wegen der im Investitionsbereich üblichen langen Lieferfristen womöglich erst im Abklingen der Konjunktur ertragswirksam werden kann. In einer solchen Situation aber ist ein hoher Kapitalstock für eine Nettoinvestition eher hinderlich als fördernd. Und analoge Überlegungen gelten auch für Erhaltungs- und Ersatzinvestitionen: In einer Phase schwacher Konjunktur steigt die Neigung, Anlagen oder ganze Firmenteile abzustoßen, sich also "gesundzuschumpfen", wenn ein großer Kapitalstock existiert. Akzeptiert man diese Überlegungen, dann hat der börsennotierte Marktwert einer Firma nicht mehr die Rolle eines globalen "Gewinnerwartungsindikators", sondern spiegelt vor allem Gewinnerwartungen wieder, die über den allgemeinen Konjunkturverlauf hinaus spezifisch für die betreffende Firma sind; die Börse spricht dann von "Sondereinflüssen".

Eine Zusammenfassung dieser Überlegungen kann man in den Schätzergebnissen (X.7) und (X.8) widergespiegelt sehen: Die bremsende Rolle eines hohen Kapitalstocks führt zu den negativen Koeffizienten bei C_{-1} bzw. C'_{-1} , während die Rolle des börsennotierten Wertes als Indikator von neben dem allgemeinen Konjunkturverlauf firmentypischen Gewinnerwartungen zu einem Anwachsen der Koeffizienten bei F_{-1} bzw. F'_{-1} führt; diese Variablen werden also in ihrer Rolle als Gewinnindikatoren stärker betont.

Um die Stabilität des Ausgangsmodells zu prüfen, wurden für alle bei diesem Modell bisher verwendeten Schätzansätze ergänzende Rechnungen durchgeführt, jetzt aber mit zwei getrennten Zeitreihen von jeweils dem Umfang $T = 10$. Die erste Zeitreihe bildeten die Meßwerte der ersten zehn Jahre von 1935 bis 1944, die zweite Zeitreihe die Wertematrizen der letzten zehn Jahre von 1945 bis 1954¹⁾. Die Ergebnisse sind auf den Seiten A.X.6 bis A.X.9 wiedergegeben; für die Zellner-Schätzungen wurde ein Zweigleichungsmodell mit den jeweils korrespondierenden Zeitpunkten gewählt. Ein 4 - Gleichungsmodell hätte für den simultanen Zellner-Schätzer nur Sinn, wenn man z. B. über einen zehnjährigen Konjunkturzyklus neben einer Korrelation zwischen zeitgleichen Residuen der verschiedenen Gleichungen auch eine Autokorrelation innerhalb der Gleichungen mit zehnjährigem Lag annimmt. Für eine Annahme eines Lags von exakt zehn Jahren aber sprechen weder die wirtschaftliche Theorie noch ist z. B. aus den Residuen der Zellner-Schätzungen für den

1) Zur schnellen Unterscheidung in den Gleichungsausdrücken wurden die Variablen der ersten Zeitreihe von 1935 - 1944 mit "a" gekennzeichnet, die der zweiten Zeitreihe von 1945 - 1954 mit "b".

vollen Zeitreihenumfang ¹⁾ eine solche Autokorrelation zu sehen ²⁾. Hierbei ist allerdings zu berücksichtigen, daß eine solche Autokorrelation der Residuen den Annahmen des Zellner - Schätzers widerspräche ³⁾ und damit eine Schätzung unter Verwendung des vollen Zeitreihenumfangs unzulässig wäre.

Die wichtigste Gemeinsamkeit der Ergebnisse der Schätzungen in den Gleichungen (X.9) bis (X.20) ist die Instabilität der Koeffizienten gegenüber den Schätzungen mit vollem Zeitreihenumfang, wobei neben den bei C_{t-1} besonders die bei C'_{t-1} (also den Westinghouse-Werten) stehenden Koeffizienten auffallen: Sie zeigen für die Vorkriegs- und Kriegsjahre ⁴⁾ signifikant positive, für die Nachkriegsjahre aber signifikant negative Werte ⁵⁾. Daß die Investitionsentscheidungen in den beiden angesprochenen Zeiträumen vor einem veränderten wirtschaftlichen Hintergrund erfolgten, wurde bereits angesprochen ⁶⁾. Hieraus kann man den Schluß ziehen, daß also entweder durch den SUR-Ansatz nach Zellner mit jeweils zwei Exogenen in den Gleichungen dieses Verhalten unzureichend abgebildet wird oder daß tatsächlich ein Strukturbruch des Investitionsverhaltens stattgefunden hat. Diese Unterscheidung aber kann nur ein Ökonometriker nach wirtschaftstheoretischen Überlegungen treffen.

Dasselbe Umkippen des Vorzeichens der bei C_{t-1} bzw. C'_{t-1}

1) Vgl. Abbildung (X.D) und (X.E) im Anhang Seite A.X.3

2) Eine versuchsweise Schätzung mit dem iterativen Zellner-Schätzer ergab keine Konvergenz; die Schätzungen jeder Iteration oszillierten unkontrolliert

3) Vgl. Seite 52 f.

4) Also für die mit "a" gekennzeichneten Werte

5) Also für die mit "b" gekennzeichneten Werte

6) Vgl. die Anmerkungen Seite 86

stehenden Schützkoeffizienten zeigt sich in noch stärkerem Maße in den Ergebnissen (X.21) bis (X.24) nach dem Schätzansatz (42). Hier sind die gleichen Entscheidungen wie bei dem Zellner - Schätzer zu fällen, wobei allerdings noch ein zusätzliches Untersuchungskriterium hinzukommt. Auch hier ist also zu fragen, ob das vorliegende Modell zweier nur durch einen gemeinsamen Ausschluß eines Regressors korrelierter Gleichungen mit jeweils zwei spezifizierten Regressoren das Investitionsverhalten unzureichend abbildet und deshalb instabil ist oder ob tatsächlich ein Strukturbruch der Investitionsdeterminanten stattgefunden hat. Darüberhinaus sind jetzt aber Überlegungen anzustellen, ob die dem Zellner-Schätzer und dem Ansatz nach (42) gemeinsame Inkonsistenz für den Ansatz (42) durch einen Übergang zu dem konsistenten Ansatz (49) beseitigt werden kann, ob man also für die endogenen Proxies geeignete Instrumentalvariablen finden kann. Hierfür aber bietet die Untersuchung von Boot und de Witt die Auswahl von acht weiteren Firmen ¹⁾. Unter diesen befindet sich leider keine weitere Firma aus der Elektrobranche wie die beiden schon eingeführten Firmen ²⁾, was die Wahl einer geeigneten Instrumentalvariablen erleichtert hätte. Für die schließlich ausgewählte Firma US Steel sprechen nach Meinung des Verfassers zwei Gründe. Erstens wird, wie schon ³⁾ dargelegt wurde, der ausgeschlossene Regressor im makroökonomischen Bereich globaler Konjunkturdaten vermutet; gerade die Stahlindustrie aber ist stark von der allgemeinen Konjunktur abhängig. Den zweiten Grund liefern statistische

1) Dies sind: General Motors Co., US Steel Co., Chrysler Co., Atlantic Refining, IBM, Union Oil, Goodyear, Diamond Match.

2) Im untersuchten Zeitraum war IBM noch nicht der Elektro- bzw. Elektronikkonzern wie heute, sondern war hauptsächlich im mechanischen Bürobereich tätig .

3) Vgl. Seite 89 f.

Gesichtspunkte: Je stärker eine Endogene von dem ausgeschlossenen Regressor abhängig war, desto stärker wird sich dies auf den Durbin-Watson-Koeffizienten als Fehlspezifikationsindikator ¹⁾ auswirken. Je stärker aber auch die Residuen der unterspezifizierten Regressionsschätzungen durch die Orthogonalanteile des ausgeschlossenen Regressors geprägt werden ²⁾, ein desto höherer Korrelationskoeffizient dieser Residuen zwischen den Gleichungen ist zu erwarten, wenn ein gemeinsamer ausgeschlossener Regressor existiert. Diese beiden Indikatoren sind bei der Investitionsgleichung von US Steel am deutlichsten vereinigt mit einem Durbin-Watson-Koeffizient von 1.09 und Residuenkorrelationen $r(\underline{v}_s \underline{v}_g) = .43$ und $r(\underline{v}_s \underline{v}_w) = .61$ ³⁾. Neben den aus diesem 3 - Gleichungsmodell erhaltenen OLS- und Zellnerschätzungen des Anhangs Seite A.X.10 und A.X.11 sind die jetzt besonders interessierenden Ergebnisse der Schätzungen nach Ansatz (49) auf Seite A.X.12 wiedergegeben in den Gleichungen (X.34) bis (X.37). Diese Schätzungen stimmen im wesentlichen mit den Ergebnissen nach Ansatz (42) überein ⁴⁾, was (bei sonst richtigen Modellannahmen !) auf eine geringe Verzerrung dieses Ansatzes (42) schließen läßt. Analog zu dem Vorgehen bei dem Beispiel A.VI. Seite 77 ⁵⁾ wurden auch hier neben der als Instrumentalvariablen dienenden Endogenen I_g (Investition von US Steel) auch unterschiedlich große Teilmengen der

1) Vgl. Seite 87 oder Schneeweiß (1974) S. 188 f.

2) Vgl. Seite 56 ff.

3) Vgl. Boot & de Witt (1960) S. 9 und S. 23 .
Um eine Anhäufung von "f"-Indizierungen zu vermeiden, wurden die zu US Steel, Westinghouse und General Electric gehörenden Variablen mit $\underline{s}$, $\underline{w}$ und $\underline{g}$ gekennzeichnet.

4) Diese Ergebnisse sind im Anhang Seite A.X.4 in den Gleichungen (X.7) und (X.8)

5) Vgl. Anhang Seite A.VI.2

Regressoren von I_s , also F_s und C_s , als Instrumentalvariablen verwendet, um einen Effizienzvergleich durchführen zu können. Hier zeigte sich, daß die Schätzkoeffizienten von dem Umfang des Instrumentalvariablensatzes kaum beeinflusst wurden. Anzumerken ist, daß das teilweise Absinken von s^2 bei kleiner werdendem Instrumentalvariablensatz keinen Rückschluß auf die Effizienz eines Schätzers zuläßt ¹⁾.

Zur Ergänzung durchgeführte Stabilitätsuntersuchungen bei zeitlich halbiertem Datensatz sind im Anhang Seite A.X.13 für $t = 1 - 10$ und auf Seite A.X.14 wiedergegeben; sie zeigen in Übereinstimmung mit den anderen Schätzansätzen denselben "Strukturbruch" für die bei der Variablen C stehenden Koeffizienten. Auch hier kann eine Bewertung nur aus wirtschaftstheoretischer Sicht erfolgen, denn selbst die sehr hohen adjustierten Bestimmtheitsmaße der Schätzergebnisse spiegeln ja hauptsächlich den für Zeitreihenanalysen typischen Einfluß des gemeinsamen Trends der Daten wieder und können ohne weiteres theoretisch nicht fundierte Nonsens-Korrelationen sein.

Faßt man die theoretischen und praktischen Ergebnisse dieses Kapitels zusammen, so läßt sich festhalten:

- a) Für "seemingly unrelated regression" - Modelle lassen sich unter bestimmten Modellvoraussetzungen konsistente Schätzungen sowohl für den Fall "Fehler in den Regressoren" wie auch für den Fall "gemeinsames Fehlen von Regressoren" konstruieren. Wesentliche Voraussetzung ist die wechselseitige Unkorreliertheit der Gleichungsresiduen des "wahren" Modells; alle Ansätze gehen davon aus, daß für

1) So gilt ja z. B. für den gegenüber OLS asymptotisch effizienten Zellner-Schätzer $s^2_{\text{Zellner}} \approx s^2_{\text{OLS}}$.

eine nichtverschwindende Kovarianz zwischen den Residuen der Einzelgleichungen des Systems eine von der (wirtschaftlichen) Theorie abgeleitete Begründung im gemeinsamen Auftreten von Proxies oder im gemeinsamen Ausschluß von Regressoren vorliegt und zu einem entsprechenden Schätzansatz führt.

- b) An simulierten Modellen gewonnene Ergebnisse zeigen, daß der bereits bekannte Zellner - Schätzer vor allem in seiner iterierenden Version bei Modellen des unter a) beschriebenen Typs trotz seiner Inkonsistenz zu erstaunlich guten Ergebnissen führen kann. Sein Mangel besteht nach Meinung des Verfassers neben einem unter Umständen großen Rechenaufwand für die Iterationen vor allem im Fehlen einer theoretischen Absicherung, ob und zu welchen Grenzwerten diese Iterationen konvergieren, wenn ein unter a) beschriebenes Modell vorliegt.
- c) Die sinnvolle, d. h. modelladäquate Anwendung der neuen Schätzansätze steht und fällt mit den ihnen zugrundeliegenden Modellannahmen, die a priori vor dem Schätzansatz von dem Ökonometriker formuliert werden müssen. Vor allem die auf den Fall gemeinsam ausgeschlossener Regressoren zugeschnittenen Schätzansätze dürfen auf keinen Fall zu einem versuchsweisen "Durchspielen" beliebiger Spezifikationen verleiten; ein solches Vorgehen würde nur zu einem nicht fundierten Sammeln von "besten" Nonsens-Korrelationen führen und die Statistik ihrer Aufgabe als Testinstrument vorformulierter Hypothesen entkleiden.

§ 7. Die Bedeutung der Struktur- und reduzierten Form des
 interdependenten ökonomischen Gleichungssystems

Um in den folgenden Ausführungen bestimmte Größen simultaner ökonomischer Gleichungssysteme in eindeutiger Kurzschreibweise ansprechen zu können, soll folgende Symbolik eingeführt werden:

Ein simultanes Gleichungssystem bestehe aus G Einzelgleichungen, an denen genau G endogene Variable y_i ($i=1, \dots, G$) und K exogene ¹⁾ Variable x_j ($j=1, \dots, K$) beteiligt sind. Von jeder endogenen und exogenen Variablen liegen T Meßwerte vor. Jede Endogene y_i sei in einer Regressionsbeziehung als abhängige Variable aller Exogenen darstellbar; der durch diese Beziehung nicht erklärte, nicht meßbare Rest sei mit v_i ($i=1, \dots, G$) bezeichnet. Alle endogenen und exogenen Variablen y_i und x_j seien durch Spaltenvektoren der Länge T dargestellt, ebenso die Größen v_i . Damit kann für jede Endogene die Regressionsbeziehung

$$\begin{aligned} y_{i1} &= x_{11}\pi_{i1} + x_{21}\pi_{i2} + \dots + x_{j1}\pi_{ij} + \dots + x_{K1}\pi_{iK} + v_{i1} \\ &\vdots \\ y_{it} &= x_{1t}\pi_{i1} + x_{2t}\pi_{i2} + \dots + x_{jt}\pi_{ij} + \dots + x_{Kt}\pi_{iK} + v_{it} \\ &\vdots \\ y_{iT} &= x_{1T}\pi_{i1} + x_{2T}\pi_{i2} + \dots + x_{jT}\pi_{ij} + \dots + x_{KT}\pi_{iK} + v_{iT} \end{aligned}$$

oder kurz $y_i = X\pi_i + v_i$ geschrieben werden.

Diese Darstellung läßt sich zur reduzierten Form für alle endogenen Variablen zusammenfassen zu

$$(y_1; y_2; \dots; y_i; \dots; y_G) = (X\pi_1; X\pi_2; \dots; X\pi_i; \dots; X\pi_G) + (v_1; v_2; \dots; v_i; \dots; v_G)$$

oder kurz $Y = X\Pi + V$. Diese Matrizen haben die Dimensionen

$$\begin{array}{c} G \\ \boxed{Y} \end{array} = \begin{array}{c} T \\ \boxed{X} \end{array} \begin{array}{c} K \\ \boxed{\Pi} \end{array} + \begin{array}{c} G \\ \boxed{V} \end{array}$$

1) Unter "exogenen" seien exogene und vorherbestimmte Variablen zusammengefaßt

Dabei werden folgende Annahmen getroffen:

- 1) $E(v_{it}) = 0$ 2) $E(v_{it})^2 = E(v_i)^2 = \sigma^2(v_i)$
- 3) $E(v_{it}v_{i\tau}) = 0$ für $t \neq \tau$ 4) $E(x'_{j-i}v_i) = 0$ für alle i und j
- 5) $r_B(X) = K$ 6) $r_B(I) = G$

Für dieses System existiere eine interdependente Darstellung in der Strukturform, die für eine beliebige i -te Gleichung neben der i -ten Endogenen noch weitere $g_i - 1$ Endogene und k_i Exogene beinhalte nach

$$Y\beta_i = X\gamma_i + u_i \quad \text{mit} \quad u_i = Y\beta_i - X\gamma_i = v\beta_i$$

Diese Matrizen haben die Dimensionen

$$\begin{array}{c} G \\ T \end{array} \begin{array}{|c|} \hline Y \\ \hline \end{array} \begin{array}{c} 1 \\ G \end{array} \begin{array}{|c|} \hline \beta_i \\ \hline \end{array} = \begin{array}{c} K \\ T \end{array} \begin{array}{|c|} \hline X \\ \hline \end{array} \begin{array}{c} 1 \\ K \end{array} \begin{array}{|c|} \hline \gamma_i \\ \hline \end{array} + \begin{array}{c} 1 \\ T \end{array} \begin{array}{|c|} \hline u_i \\ \hline \end{array}$$

Man kann jede Strukturgleichung normieren, z. B. so, daß $\beta_{ii} = 1$ ist. Mit allen G Gleichungen kann man die gesamte Strukturform des Systems darstellen nach

$$Y(\beta_1; \dots; \beta_1; \dots; \beta_G) = X(\gamma_1; \dots; \gamma_1; \dots; \gamma_G) + (u_1; \dots; u_1; \dots; u_G)$$

oder kurz $YB = X\Gamma + U$ mit $U = VB$

Diese Matrizen besitzen die Dimensionen

$$\begin{array}{c} G \\ T \end{array} \begin{array}{|c|} \hline Y \\ \hline \end{array} \begin{array}{c} G \\ G \end{array} \begin{array}{|c|} \hline B \\ \hline \end{array} = \begin{array}{c} K \\ T \end{array} \begin{array}{|c|} \hline X \\ \hline \end{array} \begin{array}{c} G \\ K \end{array} \begin{array}{|c|} \hline \Gamma \\ \hline \end{array} + \begin{array}{c} G \\ T \end{array} \begin{array}{|c|} \hline U \\ \hline \end{array}$$

Dabei werde angenommen, daß B den vollen Rang habe. Zwischen Struktur- und reduzierter Form eines Gleichungssystems besteht die Beziehung $YBB^{-1} = X\Gamma B^{-1} + UB^{-1} = YI = X\Pi + V$ mit $\Pi = \Gamma B^{-1}$ und $V = UB^{-1}$ bzw. $U = VB$.

Die Anzahl der Elemente des Vektors β_i , die ungleich Null sind, werde mit r_i bezeichnet; die Anzahl der Elemente des Vektors γ_i , die ungleich Null sind, werde mit k_i bezeichnet. Dann besagt

1) Vgl. z. B. Schönfeld (1971) Band II S. 206

7.1. Ein erläuterndes Beispiel

Für die richtige Interpretation und daraus abgeleitete Schlußfolgerungen interdependenter ökonomischer Gleichungssysteme ist es von zentraler Bedeutung, das Wesen eines mit "klassischen" ¹⁾ Schätzmethoden geschätzten interdependenten Gleichungssystems zu untersuchen. Dies kann am deutlichsten an einem kleinen Beispiel geschehen.

Gegeben sei ein Datensatz mit $T \rightarrow \infty$ Messungen für 3 endogene und 3 exogene Variablen ²⁾, der für eine Schätzung mit einem "klassischen" Schätzverfahren folgende Strukturform ergab:

$$1 \cdot \underline{y}_1 + 1 \cdot \underline{y}_2 = -1 \cdot \underline{x}_1 + \underline{u}_1 \quad (1s)$$

$$1 \cdot \underline{y}_2 + 1 \cdot \underline{y}_3 = -1 \cdot \underline{x}_2 + \underline{u}_2 \quad (2s)$$

$$1 \cdot \underline{y}_1 + 1 \cdot \underline{y}_3 = -1 \cdot \underline{x}_3 + \underline{u}_3 \quad (3s)$$

Die zugehörige reduzierte Form lautet

$$\underline{y}_1 = -\frac{1}{2} \underline{x}_1 + \frac{1}{2} \underline{x}_2 - \frac{1}{2} \underline{x}_3 + \underline{v}_1 \quad (1r)$$

$$\underline{y}_2 = -\frac{1}{2} \underline{x}_1 - \frac{1}{2} \underline{x}_2 + \frac{1}{2} \underline{x}_3 + \underline{v}_2 \quad (2r)$$

$$\underline{y}_3 = +\frac{1}{2} \underline{x}_1 - \frac{1}{2} \underline{x}_2 - \frac{1}{2} \underline{x}_3 + \underline{v}_3 \quad (3r)$$

wobei die Bedingungen erfüllt seien, die in der Einführung dieses Kapitels angegeben wurden.

Derselbe Datensatz läßt aber für die einzelnen Strukturgleichungen andere Darstellungsweisen zu, von denen nur diejenigen aufgeführt werden sollen, in denen die Variable $\underline{y}_1$ enthalten ist und die nicht zu einem Widerspruch zu den Koeffizienten der reduzierten Form führen ³⁾. Dies gilt für die folgenden 5 Strukturgleichungen:

$$I. \quad \underline{y}_1 + \underline{y}_2 = -\underline{x}_1 + \underline{u}_1^{(I)} : \hat{=} \quad (1s)$$

$$II. \quad \underline{y}_1 + \underline{y}_3 = -\underline{x}_3 + \underline{u}_1^{(II)} : \hat{=} \quad (3s)$$

1) Unter "klassisch" werden hier alle konsistenten Schätzverfahren subsummiert, für die $\text{plim}(\hat{\beta}) = \beta$ gilt. Eine Begründung für diese Bezeichnung wird im nächsten Kapitel gegeben.

2) Im Beispiel A.XI. des Anhangs ist T natürlich endlich; die Bedingung $\underline{x}'\underline{v} = 0$ ist aber für alle i und j erfüllt.

3) Diese "Darstellungsweisen" können also mit allen Schätzverfahren für über- und exakt identifizierte Gleichungen konsistent geschätzt werden.

$$\text{III. } \underline{y}_1 - \underline{y}_2 = \underline{x}_2 - \underline{x}_3 + \underline{u}_1^{(\text{III})} : \hat{=} (1r) - (2r)$$

$$\text{IV. } \underline{y}_1 - \underline{y}_3 = -\underline{x}_1 + \underline{x}_2 + \underline{u}_1^{(\text{IV})} : \hat{=} (1r) - (3r)$$

$$\text{V. } \underline{y}_1 = -\frac{1}{2}\underline{x}_1 + \frac{1}{2}\underline{x}_2 - \frac{1}{2}\underline{x}_3 + \underline{u}_1^{(\text{V})} : \hat{=} (1r)$$

Die Varianz von $\underline{u}_1^{(\text{V})} = \underline{y}_1$ kann dann unverzerrt geschätzt werden nach $s_{(\text{V})}^2 = \sum_{t=1}^T v_{1t}^2 / (T-3)$. Die Varianzen der übrigen 4 Alternativen können konsistent geschätzt werden¹⁾ wegen

$$U = VB \text{ bzw. } \underline{u}_i = V\underline{\beta}_i \text{ zu}$$

$$s_{(\text{I})}^2 = \sum (v_{1t} + v_{2t})^2 / (T-2) ; \quad s_{(\text{II})}^2 = \sum (v_{1t} + v_{3t})^2 / (T-2)$$

$$s_{(\text{III})}^2 = \sum (v_{1t} - v_{2t})^2 / (T-3) ; \quad s_{(\text{IV})}^2 = \sum (v_{1t} - v_{3t})^2 / (T-3)$$

Zu untersuchen ist nun, nach welchen Kriterien ein Ökonometriker, der durch die von ihm ausgewählten endogenen und exogenen Variablen ein solches Bündel von Alternativen erhalten hat, unter den Gleichungen I bis V wählen kann und ob eine solche Wahl überhaupt zulässig und sinnvoll ist.

Die Gleichung V, die unverzerrt und konsistent mit OLS geschätzt werden kann, kann mit den bekannten statistischen Tests auf Signifikanz der Koeffizienten untersucht werden; wenn die Spezifikation der exogenen Variablen dem "wahren" Modell entspricht, liefert jede Spezifikation mit einem abweichenden Satz von Regressoren ein größeres s^2 wegen des Kriteriums von Theil²⁾. Nehmen wir an, daß die durch das Modell vorgegebene Wahl der Exogenen dem "wahren" Modell entspricht und die Annahmen 1) bis 5) der Einleitung dieses Kapitels erfüllt sind; dies gelte auch für die Gleichungen (2r) und (3r).

Dann gilt - wegen $T \rightarrow \infty$ - $E(\underline{y}_i) = \underline{\eta}_i = X\hat{\underline{\pi}}_i$. Alle diese Regressionsschätzungen der reduzierten Form können dann in der bekannten Weise interpretiert werden: Der Erwartungswert von $\underline{y}_1$ ändert sich z. B. um den Betrag $\pi_{13} = -\frac{1}{2}$, wenn sich

1) Auf die mögliche unterschiedliche Wahl des Nenners bei der Schätzung von s^2 wird später noch eingegangen.

2) Vgl. Kapitel 4

x_3 um den Betrag 1 ändert und x_1 und x_2 unverändert bleiben. y_1 wird also von x_1 , x_2 und x_3 determiniert, wobei sich dem Einfluß der drei Exogenen eine "zufällige" Störung v_1 überlagert.

Betrachtet man die reduzierte Form als eine Darstellungsform eines interdependenten Systems, die wegen der nur mit ihr möglichen Aussagen ¹⁾ auf jeden Fall neben der Strukturform dargestellt werden kann und nicht als alternativ zu dieser betrachtet werden muß, so bleibt für die Strukturform noch die "Wahl" zwischen den "Alternativen" I bis IV.

Diese Gleichungen I bis IV sind allerdings keine Alternativen, denn sie sind ja nichts anderes als äquivalente Transformationen ²⁾ des Gleichungssystems. Alle Gleichungen sind "richtig", wenn die reduzierte Form richtig spezifiziert ist, d. h. wenn für die Endogenen, die erklärt werden sollen, die "richtigen" Exogenen ausgewählt wurden. Da dies für den vorliegenden Fall angenommen wurde, sind alle vier Gleichungen nichts anderes als verschiedene Darstellungsformen des durch die reduzierte Form definierten Gleichungssystems. Alle vier Gleichungen beschreiben denselben stochastischen Prozess, der durch die reduzierte Form als Abhängigkeit jeweils einer Endogenen von mehreren Exogenen oder in der Strukturform als die Abhängigkeit mehrerer Endogener gemeinsam ³⁾ von einem Satz von Exogenen, der in der ursprünglichen Darstellung des Beispiels aus jeweils einer Exogenen besteht.

So besagt also z. B. die Gleichung I, daß die Endogenen y_1 und y_2 gemeinsam von der Exogenen x_1 beeinflusst werden:

1) Diese Aussagemöglichkeiten werden in Kapitel 8 untersucht.

2) Vgl. z. B. Schneeweiß (1974) S. 262 oder Fisher (1966) S. 23f.

3) Theil (1971) S. 430 spricht von "jointly and simultaneously"

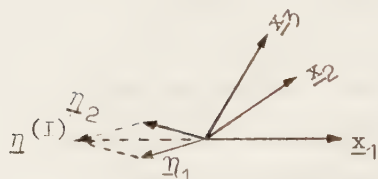
eine aus $1 \cdot \underline{y}_1$ und $1 \cdot \underline{y}_2$ bestehende endogene Variable " $\underline{y}^{(I)}$ " wird durch eine exogene Variable so beeinflusst, daß eine Veränderung von x_1 um eine Einheit diese Variable " $\underline{y}^{(I)}$ " um die Einheit $\gamma_{11} = -1$ verändert. Diese Beeinflussung von " $\underline{y}^{(I)}$ " ist von einer "zufälligen" Störung $u_1^{(I)}$ überlagert, wobei die Größe $u_1^{(I)}$ der stochastische Anteil von " $\underline{y}^{(I)}$ " ist, der von dem Satz aller Exogenen, die in dem Modell vorhanden sind, nicht erklärt werden kann. Nach dem Ansatz $u_1^{(I)} = \beta_{11}^{(I)} \underline{y}_1 + \beta_{12}^{(I)} \underline{y}_2$ kann man also die Größe $u_1^{(I)}$ als die Summe der stochastischen Elemente veranschaulichen, die die Variablen $\underline{y}_1$ und $\underline{y}_2$ in die Variable " $\underline{y}^{(I)}$ " einbringen. Und in dieser Variablen " $\underline{y}^{(I)}$ " stehen die Variablen $\underline{y}_1$ und $\underline{y}_2$ in einem Gleichgewicht, das durch das Verhältnis $\beta_{11} : \beta_{12} = 1 : 1$ beschrieben wird. Somit ist die Aussage der Gleichung I: Eine Gesamtheit von abhängigen Variablen " $\underline{y}^{(I)}$ " wird von einer unabhängigen Variablen $\underline{x}_1$ so beeinflusst, daß eine Vergrößerung dieser unabhängigen Variablen $\underline{x}_1$ um eine Einheit die abhängige Variable " $\underline{y}^{(I)}$ " um eine Einheit verringert ¹⁾, wenn diese abhängige Variable als Summe von $1 \cdot \underline{y}_1 + 1 \cdot \underline{y}_2$ gebildet wird. Wichtig bei dieser Aussage ist das Zusammenfassen der endogenen Variablen zu einer einzigen neuen Endogenen. Dieses Zusammenfassen wird auch bei der Schätzung des Vektors $\underline{\beta}$, der die Aufteilung dieser Endogenen auf die einzelnen als Summanden auftretenden endogenen Variablen $\underline{y}_1$ und $\underline{y}_2$ angibt, bei z. B. Indirect Least Squares (ILS) und Limited information single equation/Maximum Likelihood (LISE/ML, $\hat{=}$ Least Variance Ratio=LVR) deutlich ²⁾: Unter dem Satz von Endogenen, die $\underline{y}^{(I)}$ bilden, wird keine der beteiligten Endogenen $\underline{y}_1$ oder $\underline{y}_2$ ausgezeichnet, es wird für die Schätzung nicht "nach einer Endo-

1) Dieselbe Aussage kann man selbstverständlich auch aus der reduzierten Form für $\underline{y}_1$ und $\underline{y}_2$ erhalten.

2) Vgl. z. B. Wonnacott & Wonnacott (1970) Chapter 19

genen aufgelöst". Bei dem für die Schätzung von Strukturgleichungskoeffizienten wohl am häufigsten benutzten Schätzverfahren, nämlich Two Stage Least Squares (TSLS), wird aber für die Schätzung nach einer (scheinbar!) abhängigen Endogenen aufgelöst; die restlichen Endogenen der jeweiligen Strukturgleichung nehmen (scheinbar!) die Rolle von Regressoren an. Und das Schätzergebnis ist, zumindest für endliches T , abhängig von der Wahl der als "abhängig" deklarierten endogenen Variablen. Daß dies jedoch nur ein "Fehler", d. h. eine Verzerrung des Schätzverfahrens ist, die bei $T \rightarrow \infty$ verschwindet, wird nur allzuleicht übersehen und führt zu folgenreichen Fehlinterpretationen. Hierauf wird in diesem Kapitel noch eingegangen. Zur Veranschaulichung des durch Gleichung I beschriebenen Modells sollen die Abbildungen 1 und 2 dienen, die die Modellzusammenhänge im Testraum darstellen. Dabei sind, um im dreidimensionalen Raum zu bleiben, statt der Vektoren $\underline{y}_1$ und $\underline{y}_2$ deren Erwartungswerte η_1 und η_2 dargestellt.

Abb. 1



reduzierte Formen (1r);(2r)

Abb. 2



Strukturform nach Gl. I

Diese Darstellung macht deutlich, wie die beiden Vektoren η_1 und η_2 , die in der zur reduzierten Form gehörenden Abb.1 im durch die drei Exogenen des Gesamtmodells aufgespannten dreidimensionalen Raum liegen, durch die von dem Koeffizientenvektor $\underline{\beta}^{(I)}$ vorgegebene Linearkombination einen neuen Endogenenvektor $\underline{\eta}^{(1)}$ bilden, der zu seiner Darstellung nicht mehr die von $\underline{x}_2$ und $\underline{x}_3$ aufgespannte Ebene benötigt, sondern nur

noch von $\underline{x}_1$ abgebildet wird: $\underline{y}^{(I)}$ ist jetzt nur noch von $\underline{x}_1$ abhängig bzw. wird durch $\underline{x}_1$ dargestellt.

Da $\underline{y}^{(I)} = \gamma_{11}^{(I)} \underline{x}_1 + \underline{u}_1^{(I)}$ ist mit $\underline{u}_1^{(I)} = \beta_{11}^{(I)} \underline{v}_1 + \beta_{12}^{(I)} \underline{v}_2$, wobei ja sowohl $\underline{v}_1$ wie auch $\underline{v}_2$ nach den Modellannahmen

orthogonal zu $\underline{x}_1$ sind, könnte man zur Angabe des Bestimmtheitsmaßes dieser Regression eine Varianzzerlegung von $\underline{y}^{(I)}$ durchführen und damit das Bestimmtheitsmaß angeben nach

$$B = \sigma^2(\hat{\gamma}_{11}^{(I)} \underline{x}_1) / \sigma^2(\underline{y}^{(I)}) \text{ oder allgemein } B = \sigma^2(\hat{\underline{y}}^{(I)}) / \sigma^2(\underline{y}^{(I)}) \quad 1)$$

Dieses Bestimmtheitsmaß ist, wie man sofort sieht, unabhängig von irgendeiner Normierung des Vektors $\underline{\beta}$ und darf nicht mit der falschen Berechnung $B^f = \sigma^2(-\hat{\beta}_{12} \underline{v}_2 + \hat{\gamma}_{11} \underline{x}_1) / \sigma^2(\underline{v}_1)$ verwechselt werden, die ja auf einer wegen $\text{cov}(\underline{v}_2; \underline{u}_1) \neq 0$ unzulässigen Varianzzerlegung beruht ²⁾. Das Bestimmtheitsmaß B ist aber kein Maß für die Güte der Gleichung I, sondern lediglich ein Maß für den Varianzanteil der in dieser Gleichung auftretenden Exogenen, wenn man das in $\underline{y}^{(I)}$ postulierte Gleichgewicht von $\underline{v}_1$ und $\underline{v}_2$ als Modellvorstellung postuliert. Und da $\underline{v}_1$ und $\underline{v}_2$ durch ihre reduzierte Form als im dreidimensionalen $\underline{x}_1 \underline{x}_2 \underline{x}_3$ -Raum liegend definiert wurden, entfallen zwar die $\underline{x}_2$ bzw. die $\underline{x}_3$ -Anteile durch die Formulierung des Gleichgewichts $\underline{v}^{(I)} = 1 \cdot \underline{v}_1 + 1 \cdot \underline{v}_2$, aber zur fehlerfreien Abbildung dieses Gleichgewichts im $\underline{x}_1 \underline{x}_2 \underline{x}_3$ -Raum ist noch eine " $\underline{x}_1$ -Komponente" $\gamma_{11}^{(I)} \underline{x}_1$ nötig.

Zu prüfen ist noch, ob die Varianz von $\underline{u}_1^{(I)}$ ein Maß für die Güte der Gleichung I ist. Nach den bisherigen Ausführungen ist es sofort ersichtlich, daß diese Varianz auf jeden Fall zuerst mit der Varianz von $\underline{y}^{(I)}$ standardisiert werden müßte, z. B. nach $\sigma^2(\tilde{\underline{u}}_1^{(I)}) = \sigma^2(\underline{u}_1^{(I)}) / \sigma^2(\underline{y}^{(I)})$, da ja jede

1) $\hat{\gamma}$ ist also die Regressionsschätzung von $\underline{y}$ auf $\underline{x}$

2) Zur Unzulässigkeit von B^f vgl. z. B. Christ (1966) S. 549
Ein schönes Gegenbeispiel ist auch Gl. (XI.3) im Anhang:
Die Summe der $\underline{u}_{2t}^2$ ist größer als die Summe der $\underline{y}_{2t}^2$!

Normierung des β -Vektors willkürlich ist ¹⁾. Doch selbst nach dieser Standardisierung ist $\sigma^2(\tilde{u}^{(I)})$ kein Maß für die Güte irgendeiner Abbildung, wie man dies aus der multiplen Regressionsregression kennt. Denn, wie Wold ²⁾ schreibt: "The difference between the models is a matter of the hypothetical mechanism, that is assumed to generate the empirical data... . It is not a matter of statistical estimation Nor is the difference a matter of realism, of goodness of fit between the model and the empirical data, for ... the models are different, and yet they define the same stochastic process, with the result, that they yield precisely the same goodness of fit to a given set of data." Denn $\sigma^2(\hat{u}^{(I)})$ bezeichnet ja nichts anderes als die Varianz der stochastischen Anteile von $y^{(I)}$, also der durch die vom Satz aller Exogenen des Modells nicht erklärten Anteile von y_1 und y_2 , wenn man y_1 und y_2 als Gleichgewicht darstellen will. Ob man aber dieses Gleichgewicht untersuchen will oder z. B. das durch Gleichung IV definierte, hängt von der Fragestellung des Ökonometrikers ab, nicht aber von dem durch die reduzierte Form definierten Prozeß. Wenn also ein Volkswirt z. B. die Gleichgewichtspreisbildung untersucht ³⁾, das heißt die Gleichgewichtsbeziehungen zwischen Preis (das wäre in unserem Beispiel y_1) und Menge (entsprechend y_2), dann ist ihm z. B. mit der Aussage "In einem nach Gleichung IV definierten Gleichgewicht mit y_3 (das könnte z. B. das Volkseinkommen sein) ist die Varianz der stochastischen Anteile geringer und damit sind die Koeffizienten jenes Gleichgewichts genauer (mit einer kleineren Varianz der Koeffizienten) bestimmbar" keine Antwort auf seine Frage gegeben.

1) Also z. B. $\beta_{11}=1$ oder $\beta_{12}=1$; vgl. auch Wold (1960) S.448

2) Wold (1960) S. 455

3) Z. B. bei Neumann (1973) S. 67 ff.

Man kann also zusammenfassen, daß die Varianz des stochastischen Prozesses, der der Abbildung der Endogenen durch einen Satz von Exogenen überlagert ist, durch die Wahl der Endogenen und des Satzes aller Exogenen definiert ist, also durch die reduzierte Form bestimmt wird. Diese reduzierte Form läßt sich durch äquivalente Transformation in eine Vielzahl von Strukturformen umformulieren, ohne daß ein Widerspruch zur reduzierten Form entsteht. Jede dieser Umformungen spiegelt eine unterschiedliche Modellvorstellung über das Gleichgewicht der in der jeweiligen Gleichung auftretenden Endogenen wider, aber weder die Varianz des Residuums u_i noch Signifikanztests der Schätzkoeffizienten sagen etwas über die "Güte", also die Realitätsnähe der jeweils gewählten Strukturdarstellung aus, sondern geben nur Auskunft über die Genauigkeit, mit der man die Koeffizienten der ausgewählten Gleichgewichtsdarstellung schätzen kann.

In der Praxis hat es der Ökonometriker aber nicht mit unendlich langen Zeitreihen zu tun. Die oben durchgeführten Überlegungen sind also auch für endliche T anzustellen. Da jetzt also $\hat{\Pi} \neq \Pi$ sein kann (und im allgemeinen wohl auch sein wird), muß für die folgenden Ausführungen zwischen exakter Identifikation und Überidentifikation unterschieden werden.

7.1.1. Exakte Identifikation

Für alle exakt identifizierten Strukturgleichungen gilt das bisher Gesagte unverändert: Wegen der eindeutigen Beziehung zwischen der Strukturgleichung und den zugehörigen Gleichungen der reduzierten Form ist jede exakt identifizierte Strukturgleichung eine äquivalente Transformation der reduzierten Form. Wegen der Nichtlinearität dieser Transformation kann allerdings der Koeffizientenvektor einer solchen Struktur-

gleichung lediglich konsistent geschätzt werden ¹⁾, während die Koeffizienten der reduzierten Form unverzerrt geschätzt werden können ²⁾. Damit lassen sich die Ergebnisse des vorhergehenden Abschnitts für den hier betrachteten Fall zusammenfassen: Die Schätzung der reduzierten Form definiert mit einer unverzerrten Schätzung den stochastischen Prozeß, der bei einer Abbildung von einem Satz von Endogenen auf einen Satz von Exogenen diese Abbildung überlagert nach $\underline{y}_1 = X \underline{\pi}_1 + \underline{v}_1$, wobei wegen dem endlichen Stichprobenumfang T allerdings $\hat{\pi}_1 \neq \pi_1$ sein kann. Jede exakt identifizierte Strukturgleichung, die ja in eindeutiger Weise aus dieser reduzierten Form durch äquivalente Transformation gewonnen werden kann, hat den gleichen Informationsgehalt wie die zugehörigen Gleichungen der reduzierten Form und die gleiche "Güte", d. h. Realitätsnähe. Sie ist lediglich eine andere Darstellungsweise.

7.1.2. Überidentifikation

Wenn ein Modell richtig spezifiziert ist, d. h. wenn zu einem Satz von Endogenen für die Bestimmung der reduzierten Form alle Exogenen verwendet werden, die zur Erfüllung der für eine unverzerrte Schätzung notwendigen Bedingung $E(X' \underline{v}_1) = 0$ benötigt werden ³⁾, liefert eine direkte Schätzung der reduzierten Form mit OLS unverzerrte Schätzungen ²⁾ für die Matrix Π . Da aber, wie oben erwähnt, $\hat{\Pi} \neq \Pi$ sein kann, sind bei einer nach dem Abzählkriterium überidentifizierten Strukturgleichung Widersprüche zwischen dieser Strukturgleichung und den zugehörigen Gleichungen der

1) Mit allen Schätzverfahren der k - und Doppel- k -Klasse, für die k bei exakter Identifikation den Wert 1 annimmt; Vgl. z. B. Schneeweiß (1974) S. 316

2) Bzw. konsistent, wenn im Modell stochastische Regressoren auftreten; Vgl. Kapitel 1

3) Das heißt, eventuell nicht einbezogene Regressoren müssen orthogonal zu X , dem Satz der einbezogenen Regressoren sein. Vgl. Kapitel 2

reduzierten Form möglich, die erst für $T \rightarrow \infty$ verschwinden ¹⁾. Somit bedeutet eine durch den spezifizierenden Ökonometriker vorgenommene Nullrestriktion $\gamma_{ij} = 0$ ($j = k_i + 1, \dots, K$), daß er Werte, die zwar bei unendlich großem T zu Null werden, nicht aber bei endlichem T wegen einer Abweichung der Schätzung von dem wahren Wert γ , aufgrund seines Wissens in Vorwegnahme der bei unendlichem T zu erwartenden Werte gleich Null setzt. Diese "a priori"-Restriktion bedeutet also das Einbringen von zusätzlichem Wissen über den stochastischen Prozeß in den Modellansatz, wodurch die bei endlichem T vom wahren Wert abweichenden Schätzungen für γ "korrigiert" werden. In statistischer Formulierung heißt dies, daß die Varianz der Schätzung von γ bei endlichem T noch nicht Null geworden ist und damit auch die zur Strukturform führenden äquivalenten Transformationen für die Koeffizienten der Strukturform eine Varianz aufweisen, die größer als Null ist. Da der Ökonometriker aber sicher "weiß", daß bestimmte Koeffizienten der Strukturgleichung den Wert Null annehmen müssen, die "Varianz seines Wissens" über diese Koeffizienten also Null ist, setzt er statt der mit einer endlichen Varianz behafteten, aus der reduzierten Form abgeleiteten Koeffizienten der Strukturform die Erwartungswerte für diese Koeffizienten ein. Damit aber wird deutlich, daß dieses Vorgehen nur zulässig ist, wenn er sicher ist, daß diese Erwartungswerte Null sind, denn nur dann kann er ja den aus den empirischen Daten gewonnenen direkten Schätzungen der reduzierten Form "widersprechen".

Damit existieren zwei sich widersprechende Schätzungen für die reduzierte Form: Die eine wurde durch direktes Anwenden von

1) Das Verschwinden dieser Widersprüche ist aber lediglich eine notwendige und keine hinreichende Bedingung für $\hat{\Pi} = \hat{\Pi}!$

OLS für jede der Endogenen als Regressand mit dem vollständigen Satz aller Exogenen als Regressoren (also "LSNR"¹⁾) gewonnen, die andere wurde aus der mit einem beliebigen konsistenten Verfahren geschätzten Strukturform gewonnen nach $\hat{\Pi} = \hat{\Gamma} \hat{B}^{-1}$. Damit wird dem schätzenden Ökonometriker aber der Hinweis gegeben, daß er durch die Dominanz seines "Wissens" die empirischen Daten als "falsch" abgestempelt hat. Falsch insofern, als die Meßwerte die Bedingung $X'v_i = 0$ für alle i wohl aufgrund der Modellannahmen im Erwartungswert, nicht aber für den gegebenen Datensatz mit endlichem T erfüllen. Damit aber sind natürlich auch alle auf diesem Datensatz beruhenden statistischen Schätzwerte wie z. B. $\hat{v}_i$, $\hat{u}_i$ sowie die von X und $\hat{v}$ abhängigen Schätzungen der Varianz-Kovarianz-Matrix für die $\hat{\beta}_i$ und $\hat{\gamma}_i$ "falsch" ²⁾. Der Ökonometriker muß sich also darüber im klaren sein, daß seine Formulierung des Gleichgewichts in der Strukturform und die nach der Schätzung dieser Strukturgleichungskoeffizienten daraus abgeleitete reduzierte Form mindestens in dem Maße keine richtig quantifizierte Definition des stochastischen Prozesses liefert, in dem die nach $\hat{\Pi} = \hat{\Gamma} \hat{B}^{-1}$ gewonnene Schätzung von Π im Widerspruch zu der mit LSNR gewonnenen steht.

In der Praxis steht der Ökonometriker allerdings oft vor dem Problem, daß er nicht sicher ist, ob eine Nullrestriktion richtig ist oder nicht: er vermutet sie nur. Dann könnte er in die Versuchung geraten, alternative Struktur-spezifikationen aufzustellen und unter diesen "Versuchs-

1) Die Abkürzung LSNR = Least Squares with No Restrictions ist im angelsächsischen Schrifttum üblich und wird deshalb auch im weiteren Text verwendet. Vgl. z. B. Johnston (1972) S. 418

2) Trotzdem werden z. B. bei TSLS, LISE/ML oder Three Stage Least Squares die $\hat{u}_i$ aus LSNR als Instrumentalvariablen beibehalten. Diesen Mangel versuchen iterative Instrumentalvariablenschätzer abzuschwächen; vgl. z. B. Lyttkens (1974).

Hypothesen" nach irgendwelchen Kriterien eine Auswahl zu treffen ¹⁾.

Ein nichtstatistisches Kriterium ist die Prüfung der Vorzeichen der geschätzten Koeffizienten der alternativen Strukturgleichungen aufgrund eines a-priori-Wissens des Ökonometrikers, der für einen von ihm ausgewählten Satz von Endogenen und Exogenen in einer Gleichgewichtsdarstellung bestimmte Vorzeichen der Strukturgleichungskoeffizienten erwartet. Wenn nun eine überidentifizierte Gleichung von seinen Erwartungen abweichende Vorzeichen liefert, kann dies zwei Gründe haben: Entweder war seine a-priori-Restriktion falsch, d. h. er hat durch seine Restriktion die empirischen Daten unzulässig "korrigiert", oder aber die Meßwerte verletzen die Bedingung $X'v_i = 0$ in so starkem Maße, daß die Vorzeichen der mit diesen Daten gewonnenen Schätzungen 'umkippen'. Denn man darf ja nicht vergessen, daß seine Schätzungen der Strukturgleichungskoeffizienten auf der Verwendung gerade jener Exogenen beruhen ²⁾, die er durch seine überidentifizierende Restriktion als "falsch", d. h. als mit den Fehlern v_i korreliert, abqualifiziert hat. Und im Fall mehrerer mit dem Residuum korrelierter Regressoren kann ja auch das Vorzeichen einer Regressions-schätzung 'umkippen' ³⁾. Der Ökonometriker kann also aus den seinen Erwartungen widersprechenden Vorzeichen der Strukturgleichungsschätzungen die Folgerungen ziehen, daß entweder seine a-priori-Restriktionen (also seine 'Versuchshypothese') falsch sind oder daß die empirischen Daten für eine Schätzung noch (wegen zu kleinem T) nicht

1) Es sei noch einmal betont, daß dies nur für endliche T möglich ist; nur hier sind Widersprüche zwischen Π aus LSNR und Π aus $\Pi = \Gamma B^{-1}$ möglich.

2) Sowohl bei ILS, TSLS, 3SLS wie bei LIE/ML spannen sie den Raum auf, in dem die Endogenen abgebildet werden.

3) Vgl. z. B. Maddala (1977) S. 302

geeignet sind oder beides. Zum Ablehnen der "Versuchshypothese" reicht ein Widerspruch bei den Vorzeichen allein also noch nicht aus.

Als statistische Überprüfung können aber die überidentifizierenden Restriktionen einem statistischen Test unterworfen werden ¹⁾. Beispiele für einen solchen Test geben Basman ²⁾ und Christ ³⁾. Verkürzt dargestellt, prüfen beide Tests den Zuwachs der Varianz von $\hat{Q}_i$, wenn man die $\hat{y}_i$ durch die überidentifizierenden Restriktionen nicht mehr in dem vollständigen Raum des gesamten Satzes der Exogenen darstellt, der für eine exakte Identifikation nötig wäre, sondern in einem durch die Restriktionen in seiner Dimension verringerten Unterraum ⁴⁾. Beide Tests haben allerdings den Nachteil, daß sie zum einen willkürlich eine aus allen Restriktionen einer Gleichung als die zu testende "überidentifizierende" auswählen ⁵⁾, zum anderen bei praktischen Anwendungen fast immer eine überidentifizierende Restriktion als "falsch" zurückwiesen ⁶⁾.

In der Praxis ist jedoch auch folgendes "statistisches" Verfahren üblich: Man schätzt, entweder aufgrund verschiedener konkurrierender Hypothesen oder auch ohne theoretischen Hintergrund, verschiedene Spezifikationen einer Strukturgleichung und wählt dann diejenige als die "richtige" aus, die wegen einer geringen "Residualvarianz" $s^2(u)$ oder wegen besonders hoher Signifikanz der Schätzkoeffizienten hervorsticht. Diesem Vorgehen liegen - entweder explizit formuliert oder unbewußt - zwei Vorstellungen zugrunde, die beide hier

1) Identifizierende Restriktionen können natürlich nicht getestet werden; vgl. z. B. Theil (1961) S. 507 Fußnote

2) Basman (1960)

3) Christ (1966) S. 540 f.

4) Vgl. Abbildung 2 dieses Kapitels

5) Vgl. Christ (1966) S. 541

6) Vgl. Christ (1966) S. 541 f. und Theil (1961) S. 507

nicht anwendbar sind:

a) Theils Theorem ¹⁾, daß der Erwartungswert von s^2 für

das wahre Modell kleiner oder gleich ist dem Erwartungswert von s_f^2 für ein falsch spezifiziertes Modell und

b) Die Vorstellung, daß eine Regression von einer Endogenen auf einen Satz von weiteren Endogenen und Exogenen vorliegt.

Daß die zweite Vorstellung falsch ist, sieht man sofort bei Schätzverfahren wie LISE/ML oder ILS ²⁾, bei denen schon im Schätzverfahren keine Endogene als Regressand ausgezeichnet ist, sondern die Regression eines Satzes von "jointly dependent" Endogenen auf einen Satz von Exogenen durchgeführt wird. Die Normierung irgendeines Elementes des β -Vektors zu 1 ist ja willkürlich und hat keinen Einfluß auf die Schätzung. Daß bei dem geradezu als Standardverfahren der Ökonometrie zu bezeichnenden zweistufigen Schätzverfahren TSLS die Schätzergebnisse von der Wahl derjenigen Endogenen abhängen, deren Element in dem Koeffizientenvektor β man zu 1 normiert, ist eine Eigenschaft des Schätzverfahrens, die für $T \rightarrow \infty$ verschwindet ³⁾. Man darf also nicht die (noch) unvollkommenen Eigenschaften des für dieses T verzerrten Schätzers mit dem verwechseln, was man nach dem Modellansatz schätzen will ⁴⁾. Diese Verwechslung wird sicher noch dadurch unterstützt, daß in manchen Lehrbüchern ⁵⁾ von "der abhängigen" Endogenen gesprochen wird und dadurch suggeriert wird, daß die anderen Endogenen derselben Gleichung Regressoren seien. Maddala ⁶⁾ bezeichnet

1) Vgl. Kapitel 4

2) ILS ist auch bei überidentifizierten Gleichungen ein konsistenter Schätzer! Vgl. z. B. Dhrymes (1970) S. 288

3) Da ja z. B. LISE/ML und TSLS asymptotisch unverzerrte Schätzer sind, fallen ihre Schätzungen für $T \rightarrow \infty$ zusammen.

4) Auf das Ziel einer solchen Schätzung wird im nächsten Kapitel noch einmal ausführlich eingegangen.

5) So z. B. bei Johnston (1972) S. 381

6) Maddala (1977) S. 235

dies als "contrary to the spirit of simultaneous equations because by definition y_1 and y_2 are jointly determined and we cannot label any single variable as the dependent variable in any equation", entschuldigt dieses Vorgehen allerdings damit, daß "in the back of our minds there are still some causal relationships" ¹⁾. Daß zusätzlich in manchen Lehrbüchern die interdependenten Systeme auf den Fall fehlerbehafteter Regressoren zurückgeführt werden ²⁾ und die Übereinstimmung des Lösungsansatzes mit Hilfe von Instrumentalvariablen betont wird, trägt sicher zu diesen Vorurteilen mit bei.

Während aber dieses Mißverständnis lediglich zu einer falschen Interpretation ³⁾ der durch eine Schätzung gewonnenen Koeffizienten führt, kann die unzulässige Anwendung des Theil'schen Theorems zu falschen Schätzungen führen. Dies soll anhand der ersten Gleichung des ursprünglichen Systems des Beispiels (also Gleichung (1s)) gezeigt werden.

Nehmen wir an, die erste Gleichung werde durch die Restriktion $\beta_{13} = \gamma_{11} = \gamma_{12} = \gamma_{13} = 0$ spezifiziert, der Schätzansatz laute also

$$\beta_{11}^{-1}y_1 + \beta_{12}^{-1}y_2 = \underline{u}_1 \quad 4)$$

β_{11}^{-1} werde zu 1 normiert und es werde mit TSLS geschätzt.

Dann gilt mit $\underline{y}_2 = X(X'X)^{-1}X'y_2$ und $X = (\underline{x}_1; \underline{x}_2; \underline{x}_3)$ ⁵⁾

$$\text{plim}(\beta_{12}^{-1}) = \beta_{12} + \frac{\text{cov}(\underline{y}_2, \underline{x}_1)}{\text{var}(\underline{y}_2)} \gamma_{11}$$

und für die zugehörige Schätzung von $s^2(\underline{u}_1)$

1) Maddala (1977) S. 236 . Vgl. hierzu auch Kapitel 8 und 9 !

2) Z. B. in dem sonst didaktisch so vorzüglichen Buch von Wonnacott & Wonnacott (1970) S. 358 ff.

3) Hierauf wird noch ausführlich eingegangen

4) Das hochgestellte Minuszeichen soll anzeigen, daß hier eine Variable (also $\underline{x}_1$) aus der Gleichung ausgeschlossen wurde

5) Die Variable $\underline{x}_1$ wurde also im Gesamtsystem beibehalten. Vgl. jedoch auch die Beispiele des Anhangs S. A.XI.3 f.

wegen

$$\begin{aligned} \underline{u}_1 &= \underline{v}_1 + \beta_{12}\underline{v}_2 + \gamma_{11}\underline{x}_1 - \frac{\text{cov}(\underline{z}_2, \underline{x}_1)}{\text{var}(\underline{z}_2)} \gamma_{11}\underline{z}_2 \\ &= \underline{v}_1 + \beta_{12}\underline{v}_2 + \gamma_{11}(\underline{x}_1 - \frac{\text{cov}(\underline{z}_2, \underline{x}_1)}{\text{var}(\underline{z}_2)} \underline{z}_2) - \frac{\text{cov}(\underline{z}_2, \underline{x}_1)}{\text{var}(\underline{z}_2)} \underline{v}_2 \\ &= \underline{v}_1 + \underline{v}_2 \underbrace{(\beta_{12} - \frac{\text{cov}(\underline{z}_2, \underline{x}_1)}{\text{var}(\underline{z}_2)} \gamma_{11})}_{\hat{= a}} + \gamma_{11}(\underline{x}_1 - \frac{\text{cov}(\underline{z}_2, \underline{x}_1)}{\text{var}(\underline{z}_2)} \underline{z}_2) \underbrace{\quad}_{\hat{= e}} \end{aligned}$$

Hierbei ist $\underline{e}$ das Residuum einer 'Hilfsregression' von $\underline{x}_1$ auf $\underline{z}_2$, also der Orthogonalanteil von $\underline{x}_1$ bezüglich $\underline{z}_2$.

Damit ist

$$\text{plim}(s_1^2) = \frac{T}{T-F} (\text{var}(\underline{v}_1) + a^2 \text{var}(\underline{v}_2) + 2a \text{cov}(\underline{v}_1, \underline{v}_2) + \gamma_{11}^2 \text{var}(\underline{e}))$$

Hierbei ist noch auf die Wahl der "Freiheitsgrade" F einzugehen. Verfolgt man die Literatur, so kann man zwei "Philosophien" ausmachen: Während z. B. Johnston ¹⁾, Maddala ²⁾ und Christ ³⁾ für F den Wert $F = g - 1 - k$ wählen, also in Analogie zur multiplen Regression formulieren und dabei die $g - 1$ Endogenen wie "Regressoren" behandeln, verwenden z. B. Schneeweiß ⁴⁾, Schönfeld ⁵⁾, Theil ⁶⁾ und Intrilligator ⁷⁾ für F den Wert Null. Für beide Alternativen kann man Gründe angeben, und vor allem Intrilligator, Maddala, Theil und Christ gehen auch ausführlich auf das Problem ein und begründen ihre Wahl: Da ja für alle Schätzverfahren bei simultanen Gleichungssystemen immer nur die Konsistenz nachgewiesen werden kann, also für $T \rightarrow \infty$, kann man für $\sigma^2(\underline{u})$ auch nur einen konsistenten und keinen unverzerrten Schätzer angeben. Und bei der Angabe eines solchen Schätzers für $\sigma^2(\underline{u})$ ist die Wahl

1) Johnston (1972) S. 384

2) Maddala (1977) S. 239, wobei vor allem die Fußnote 2 die Begründung liefert

3) Christ (1966) S. 444

4) Schneeweiß (1974) S. 329 oder S. 303

5) Schönfeld (1971) Band II S. 235, S. 237 und S. 245 f.

6) Theil (1971) S. 451, wobei die Fußnote die Begründung liefert

7) Intrilligator (1978) S. 401 Fußnote 19

eines beliebigen endlichen Wertes für F ohne Einfluß. Während aber Theil deshalb die Benutzung eines anderen Wertes von F als Null ablehnt, begründen bzw. tolerieren Intrilligator, Christ und Maddala die Wahl von $F = g - 1 - k$ als vorsichtige bzw. konservative Schätzwerte: Durch die Vergrößerung des Wertes von $s^2(\underline{u})$ durch den kleineren Nenner liegt man bei der Schätzung "auf der sicheren Seite". Beide Werte von F sind also auf ihre Weise 'vorsichtige' Wahlen: Während sich die Wahl von $F = 0$ auf die beweisbare Konsistenz beschränkt und damit auch den Anschein einer Unverzerrtheit vermeiden will, der ja möglicherweise durch die Wahl von $F = g - 1 - k$ suggeriert wird, aber nicht bewiesen werden kann, zieht die Wahl von $F = g - 1 - k$ die Sicherheit vor, $\sigma^2(\underline{u})$ nicht zu unterschätzen, wobei ja auch diese Schätzung konsistent ist.

Der Verfasser verwendet in seinen EDV-Programmen ¹⁾ für F den Wert $F = g - 1 - k$; für die Fortsetzung des gerade behandelten Problems werde aber $F = 0$ gewählt, um vor allem den Verdacht zu vermeiden, die Verkleinerung von $s^2(\underline{u}_1)$ des verwendeten Beispiels sei nur durch die Vergrößerung des Nenners verursacht. In Fortführung des Beispiels kann man dann für die Differenz $s^2(\underline{u}_1) - s^2(\underline{u}_1)$ schreiben

$$\text{plim}(s^2(\underline{u}_1) - s^2(\underline{u}_1)) =$$

$$(\beta_{12}^2 - a^2)\text{var}(\underline{v}_2) + 2(\beta_{12} - a)\text{cov}(\underline{v}_1, \underline{v}_2) - \gamma_{11}^2 \text{var}(\underline{e})$$

$$\text{oder, mit } \delta = \frac{\text{cov}(\underline{v}_2, \underline{x}_1)}{\text{var}(\underline{v}_2)}$$

$$\text{plim}(s^2(\underline{u}_1) - s^2(\underline{u}_1)) = \gamma_{11} ((2\beta_{12} - \gamma_{11}\delta)\text{var}(\underline{v}_2) + 2\text{cov}(\underline{v}_1, \underline{v}_2) - \gamma_{11}^2 \text{var}(\underline{e}))$$

$$= 2\beta_{12} \gamma_{11} \delta \text{var}(\underline{v}_2) + 2\gamma_{11} \delta \text{cov}(\underline{v}_1, \underline{v}_2) - \gamma_{11}^2 \delta^2 \text{var}(\underline{v}_2) - \gamma_{11}^2 \text{var}(\underline{e})$$

1) In den Beispielen des Anhangs sind deshalb immer die "Freiheitsgrade" ausgedruckt; eine Umrechnung für $F=0$ ist so leicht möglich.

Und dieser Ausdruck kann, in Abhängigkeit vor allem von γ_{11} und $\text{var}(\underline{e})$, durchaus positiv werden. In dem Zahlenbeispiel des Anhangs ¹⁾ sind einige krasse Beispiele durchgerechnet: Alle drei Strukturgleichungen weisen ein kleineres $s^2(\underline{u}_i)$ auf, wenn man jeweils die Exogene aus dem Strukturansatz ausschließt, dabei kippt in der ersten Gleichung (mit $\text{cov}(\underline{v}_1, \underline{v}_2) = 0$ das Vorzeichen von $\hat{\beta}_{12}$ um, in der zweiten Gleichung wird der "t"-Wert von $\hat{\beta}_{23}$ größer.

Wie man sah, kann eine unzulässige überidentifizierende Restriktion die geschätzte Varianz von $\underline{u}_i$ vermindern bzw. die Signifikanz der Strukturgleichungskoeffizienten erhöhen. Hiergegen ist man zwar für $T \rightarrow \infty$ geschützt, da es dann ja keine Überidentifikation im Sinne eines zu hohen Zeilenrangs von Π_0 gibt. Für endliche T aber ist dies ja sehr oft der Fall, und die unzulässige Anwendung von Theil's Theorem kann dann zu einem inkonsistenten Schätzer führen. Betont werden aber muß die Feststellung, daß eine Verminderung von $s^2(\underline{u})$ durch das unzulässige Ausschließen einer Variablen nur für eine Gleichung der Strukturform mit $g \geq 2$ möglich ist. Eine fehlspezifizierte Strukturgleichung ist aber, auch für $T \rightarrow \infty$, nicht mehr widerspruchsfrei aus der reduzierten Form abzuleiten; nur ein $\hat{\Pi}_0^- \neq \Pi_0$ erfüllt die zusätzliche falsche Nullrestriktion. Damit aber ist, wegen Theil's Theorem, $E(s^2(\underline{v}_1)) > E(s^2(\underline{v}_1))$. Jede Koeffizientenwahl für die Parameter π_{ij} ($j=1, \dots, K$), die nicht mit den OLS-Schätzungen aus LSNR übereinstimmt, minimiert eben das $s^2(\underline{v}_i)$ nicht, da ja die $\underline{v}_i$ keine Orthogonal-, sondern schiefwinklige Projektionen auf X sind. Diese Unterscheidung zwischen den Auswirkungen einer Fehlspezifikation auf Struktur- und reduzierte Form wird leider

1) Anhang Seite A.XI.2 Gleichungen (XI.7) bis (XI.9)

in der Literatur bisweilen vernachlässigt ¹⁾, es wird sogar zur analogen Behandlung von multipler Eingleichungsregression und interdependenter Strukturgleichungsschätzung direkt aufgefordert. So schreibt z. B. Leamer ²⁾: "The simultaneous equations model ... is a formal variant of the simple linear modell and therefore implies no interesting methodological issues that are not discussed herein". Doch genau dies trifft eben auf das von ihm ³⁾ ausführlich dargestellte Theorem von Theil nicht zu.

An dem vorgestellten Beispiel sieht man auch sofort, warum im Gegensatz zum Eingleichungsmodell im Fall interdependenter Strukturgleichungen das Theorem von Theil nicht gilt. Im Eingleichungsmodell gilt die Varianzzerlegung $\sigma^2(\underline{y}) = \sigma^2(\underline{\hat{y}}) + \sigma^2(\underline{\hat{v}})$, da ja $\underline{\hat{v}}$ mit $\underline{\hat{y}} = X \underline{\hat{\beta}}$ unkorreliert ist. So kann sich bei einer falschen Nullrestriktion - die hier ja nur Exogene betreffen kann - nur die Situation $\sigma^2(\underline{\hat{y}}^-) \neq \sigma^2(\underline{\hat{y}})$ ergeben; die Varianz $\sigma^2(\underline{\hat{v}}^-)$ ergibt sich zu $\sigma^2(\underline{\hat{v}}^-) = \sigma^2(\underline{\hat{v}}) + (\sigma^2(\underline{\hat{y}}) - \sigma^2(\underline{\hat{y}}^-))$, kann also nicht kleiner werden. Denn bei Fortlassen der Variablen $\underline{x}_1$ ändern sich zwar die Schätzkoeffizienten der restlichen Exogenen $\underline{x}_j$ ($j=1, \dots, K; j \neq 1$) soweit, daß sie den Anteil von $\gamma_1 \underline{x}_1$ erklären (d. h. 'übernehmen') können, der in dem von ihnen aufgespannten Raum liegt. Der zu $\underline{x}_j$ ($j \neq 1$) orthogonale Anteil von $\gamma_1 \underline{x}_1$ bleibt jedoch unerklärt und erniedrigt damit $\sigma^2(\underline{\hat{y}}^-)$. Eine Veränderung der Koeffizienten γ_j ($j \neq 1$) hat keinen Einfluß auf $\sigma^2(\underline{v})$, da ja die $\underline{x}_j$ orthogonal zu $\underline{v}$ sind.

Anders sieht es jedoch im Strukturgleichungsfall aus. Auch hier ändern sich die Koeffizienten der "Regressoren" $\underline{y}_i$ und $\underline{x}_j$ ($i=2, \dots, g$ bei Normierung $\beta_1=1$; $j=1, \dots, k$; $j \neq 1$), bis nur

1) Zum Beispiel bei Seibt (1968) S. 67

2) Leamer (1978) S. 3

3) Leamer (1978) S. 75

noch der zu ihnen orthogonale Anteil von $\underline{x}_1$ unerklärt bleibt. Jetzt aber ändert sich mit den Koeffizienten der Endogenen auch $\underline{u}$, da hier ja $\underline{u}$ und $\underline{y}_i$ korreliert sein können. Und dann kann, wie man an dem kleinen Beispiel sah, durch fälschliches Ausschließen eines exogenen Regressors der zusätzliche Fehler γ_{11} (Orthogonalanteil von $\underline{x}_1$ bezüglich $\underline{y}_1$ und $\underline{x}_j$) durch eine Reduzierung von $\underline{u}$ aufgrund der Änderung der β_1 aufgefangen oder sogar überkompensiert werden.

§8. Die Aussagen interdependenter Gleichungssysteme

Die in Kapitel 7 gefundenen Ergebnisse sind Anlaß, nach dem Ziel der Koeffizientenschätzung interdependenter Gleichungssysteme zu fragen. Hierbei ist vor allem die Rolle der Strukturgleichungen zu untersuchen, nachdem deutlich geworden ist, daß ihre Interpretation vollkommen von dem Interpretationschema abweicht, das man von Eingleichungssystemen gewohnt ist. Als erstes soll das "klassische" interdependente Gleichungssystem behandelt werden, das ein geschlossenes Gleichgewichtssystem darstellt.

8.1. Das geschlossene Gleichgewichtssystem

Wie schon an dem kleinen Beispiel des Kapitels 7 deutlich wurde, stellen in einem "klassischen" interdependenten System die Regressoren des Gesamtsystems diejenigen Variablen dar, die das System "determinieren": Das Verhalten einer endogenen Variablen hängt nur von dem Verhalten der sie nach der Spezifikation der reduzierten Form beeinflussenden Exogenen ab, wobei diese Beeinflussung von einer stochastischen Störung überlagert ist. Eine Beeinflussung einer Endogenen durch eine solche stochastische Störung einer anderen Endogenen findet nicht statt. Das heißt nicht, daß die Residuen aus der reduzierten Form orthogonal sein müssen; es besagt lediglich, daß eine Korrelation zwischen den Residuen der reduzierten Form der einzelnen Gleichungen des Systems weder durch den Modellansatz erklärt noch in dem Schätzverfahren verwendet wird. "Geschlossen" ist das System deshalb zu bezeichnen, weil zu einer Formulierung einer strukturellen Gleichgewichtsbeziehung nach Kapitel 7 weder zusätzliche Einflußgrößen noch weitere Wechselbeziehungen innerhalb des Systems postuliert werden müssen; Eine Strukturgleichung

beinhaltet, zumindest für $T \rightarrow \infty$, nicht mehr Informationen als die Gleichungen der reduzierten Form der beteiligten endogenen Variablen. Beide Darstellungsformen desselben Systems haben aber unterschiedliche Aussagemöglichkeiten, die im Folgenden untersucht werden.

8.1.1. Die Strukturform

Wenn bei einer empirischen Schätzung der Koeffizienten eines interdependenten Systems die Koeffizienten der Strukturform geschätzt werden, kann dies zwei Gründe haben: Entweder ist der Ökonometriker an der Strukturform selbst interessiert oder die Strukturform wird lediglich dazu benutzt, durch Einbringen von zusätzlichen überidentifizierenden Restriktionen (für endliches T !) in die Strukturform die Schätzung der reduzierten Form zu verbessern. Letzteres wird im nächsten Abschnitt behandelt.

Die Strukturform selbst ist zur statistischen Überprüfung komparativ-statischer Gleichgewichtshypothesen geeignet, d. h. sie erlaubt es, eine solche Hypothese anhand empirischer Daten zu falsifizieren. Wie die Koeffizienten einer Strukturgleichung interpretiert werden können, wurde im vorhergehenden Kapitel ausführlich besprochen und soll deshalb nicht wiederholt werden. Um den Unterschied zu den "offenen" Modellen der nächsten Abschnitte zu verdeutlichen, muß die Rolle der Residuen einer Strukturform noch untersucht werden. Wegen der Beziehung $\underline{u}_i = V\beta_i$ sieht man sofort, daß der Fehler in einer Endogenen aus der reduzierten Form keinen Einfluß auf den Erwartungswert einer Strukturgleichung hat, an der diese Endogene beteiligt ist: "Reaktionen" der Endogenen einer Strukturgleichung sind aufgrund der Modellvorstellung nur auf Änderungen der Exogenen möglich, nicht aber auf endogene

Variable bzw. auf deren stochastische Anteile. Dies ist das charakterisierende Merkmal geschlossener Gleichgewichtssysteme: Im Erwartungswert reagieren endogene Variable nur auf Veränderungen der Exogenen des Gesamtsystems, nicht aber auf irgendwelche stochastischen Störgrößen, die irgendeine endogene Variable "in das System einbringt".

8.1.2. Die reduzierte Form

Ist das Erhalten der reduzierten Form eines interdependenten Systems Haupt- oder Nebenziel der Schätzung der Koeffizienten dieses Systems, so wird die Aufgabe der reduzierten Form in der Literatur im allgemeinen durch die Bezeichnung "Prognoseform" beschrieben. Diese Bezeichnung hält der Verfasser für unglücklich, er würde die Bezeichnung "Reaktionsform" vorziehen. Denn es ist nicht einsichtig, warum hier von "prognostizieren" gesprochen wird, da ja die eigentliche "Vorhersage" bei den Exogenen stattfindet. Die endogenen Variablen sind dann lediglich eine (Linear-) Kombination der prognostizierten Exogenen. Wenn diese Exogenen also mit irgendeinem fehlerbehafteten Prognoseverfahren (z. B. Zeitreihenanalyse) bestimmt werden müssen, dann ist es schwer zu verstehen, warum nicht die Endogenen mit eben diesem Verfahren prognostiziert werden. Nur wenn die zukünftigen Werte der Exogenen fehlerfrei vorliegen (z. B. veröffentlichte Haushaltspläne, feste Absichtserklärungen o. ä.) oder wenn die Prognose der Exogenen leichter ist als die der Endogenen, wird ein Vorteil des "zweistufigen" Prognoseverfahrens (zuerst Prognose der Exogenen, dann berechnen der Reaktionen der Endogenen über $\hat{\Pi}$) ersichtlich und kann möglicherweise den bei der Bestimmung von Π auftretenden Schätzfehler überkompensieren.

Die eigentliche Aufgabe der reduzierten Form wird in der Literatur im allgemeinen unter der Bezeichnung "Wahl der optimalen Politiken" aufgeführt. Denn da die endogenen Variablen auf die exogenen Variablen (und nur auf diese!) reagieren, kann durch eine geeignete Beeinflussung dieser "Instrumente" eine gewünschte Zielgröße der Endogenen erreicht werden. Dabei erlaubt die reduzierte Form sowohl eine Quantifizierung des zur Erreichung einer Zielgröße nötigen Einsatzes der "Instrumente" ebenso wie die Kontrolle der Auswirkungen eines solchen Einsatzes auf die übrigen Endogenen des Systems: Zielkonflikte sind so quantifizierbar und können minimiert werden.

Für unendlichen Stichprobenumfang ist zur Bestimmung der Koeffizienten der reduzierten Form eine Schätzung der Koeffizienten der Strukturform nicht nötig und liefert auch keine zusätzlichen Informationen, da beide Formen ja äquivalente Darstellungen sind. Bei endlichen Stichprobenumfängen kann eine Schätzung der Strukturform aber dann eine Verbesserung der reduzierten Form nach sich ziehen, wenn in die Schätzung der Strukturform zusätzliches Wissen des Ökonometrikers in Form überidentifizierender Restriktionen eingeht. Ob diese Restriktionen in der Form des Ausschlusses exogener Variablen aus der Strukturgleichung, in Form einer geforderten Beziehung zwischen den Koeffizienten der Strukturgleichung (z. B. $\beta_{13} = \beta_{14} + \beta_{15}$) oder in Form einer Restriktion der Varianz-Kovarianz-Matrix der $\underline{u}_i$ auftritt, ist ohne Belang. Entscheidend ist lediglich, daß diese Restriktionen tatsächlich zusätzlich eingebrachtes Wissen darstellen.

8.2. Das offene "Kausalketten" - System

Das für die Identifikation einer Strukturgleichung eines interdependenten Systems notwendige Abzählkriterium ist nichts anderes als eine Formalisierung der Aussage, daß eine Gleichung eines Systems immer in einem Raum darstellbar sein muß, der durch die als einzige Einflußgrößen auftretenden exogenen Variablen und ein dazu orthogonales Störglied aufgespannt wird; Einflußgrößen außerhalb dieses Raumes gibt es nicht.

Für die weiteren Untersuchungen bedeutsamen Ausnahmefälle bilden die rekursiven Systeme, die in der angelsächsischen Literatur auch als "causal chain"-Systeme bezeichnet werden. Hier werden zur Identifikation Restriktionen an die Varianz-Kovarianz-Matrix der Residuen der Strukturform verwendet; daß diese Restriktionen über die formale Notwendigkeit hinaus auch inhaltliche Bedeutung haben, wird im folgenden gezeigt.

8.2.1. Rekursive Systeme

Ein allgemeines interdependentes System mit den Einzelgleichungen ¹⁾

$$Y_i = -Y_i \beta_i + X \gamma_i + u_i \quad (i = 1, \dots, G) \quad (1)$$

mit $Y_i = (Y_{i1}; Y_{i2}; \dots; Y_{i-1}; Y_{i+1}; \dots; Y_{G-1}; Y_G)$

ist als rekursiv definiert durch folgende zwei Bedingungen ²⁾

$$a) \quad B = \begin{pmatrix} 1 & & & \beta_{1j} \\ & 1 & & \\ & & \ddots & \\ & 0 & & 1 & 1 \end{pmatrix} \quad (2)$$

d. h. B hat eine Dreiecksgestalt mit

$$\beta_{ij} = \text{beliebig für } i > j$$

$$\beta_{ii} = 1 \quad \text{für } i = j$$

$$\beta_{ij} = 0 \quad \text{für } i < j$$

¹⁾ Die veränderte Schreibweise gegenüber Kapitel 7 hat inhaltliche Gründe und ist keine Willkür

²⁾ Vgl. z. B. Johnston (1972) S. 377 f.

b) Die Varianz-Kovarianz-Matrix der Residuen der Strukturform ist im Erwartungswert eine Diagonalmatrix, also

$$E(U'U) = \Sigma = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_G^2 \end{pmatrix} \quad (3)$$

Es besteht also keine 'mutual correlation' zwischen den einzelnen Gleichungen.

Ein solches System, das von vielen als "das eigentliche ökonometrische System" angesehen wird ¹⁾, hat eine für die Schätzung der Koeffizienten interessante Eigenschaft: Alle Koeffizienten des Strukturgleichungssystems lassen sich durch direkte Anwendung von OLS auf jede Einzelgleichung konsistent schätzen. Einen Beweis führt z. B. Johnston vor ²⁾, es soll jedoch anhand eines kleinen Modells eine von der in Lehrbüchern üblichen Form abweichende Beweisführung dargestellt werden, die nach Auffassung des Verfassers die Implikationen der Bedingung (3) besser verdeutlicht.

Gegeben sei ein Modell

$$Y_1 = X Z_1 + u_1 \quad (4.1)$$

$$Y_2 = -\beta_{21} Y_1 + X Z_2 + u_2 \quad (4.2)$$

mit den Bedingungen $E(u_{it}) = E(v_{it}) = 0$,

$$\text{plim}(\frac{1}{T} u_1' u_2) = \text{plim}(\frac{1}{T} X' u_1) = \text{plim}(\frac{1}{T} X' u_2) = 0 \quad (5)$$

Mit der Beziehung $U = VB$ ³⁾ erhält man

$$U' = \begin{pmatrix} v_1' \\ v_1' \beta_{21} + v_2' \end{pmatrix} = \begin{pmatrix} u_1' \\ u_1' \beta_{21} + u_2' \end{pmatrix} \quad (6)$$

Dann bedeutet die Forderung

$$\text{plim}(\frac{1}{T} u_1' u_2) = \text{plim}(\frac{1}{T} \text{cov}(v_1; v_1 \beta_{21} + v_2)) = 0$$

1) Vgl. z. B. die dreiteilige Ausführung über causal chains von Strotz & Wold(1960); Strotz (1960); Wold (1960)

2) Johnston (1972) S. 378

3) Vgl. Kapitel 7; Einführung der Nomenklatur

$$\beta_{21} \text{var}(\underline{v}_1) + \text{cov}(\underline{v}_1; \underline{v}_2) = 0 \quad \text{oder}$$

$$\beta_{21} = \frac{-\text{cov}(\underline{v}_1; \underline{v}_2)}{\text{var}(\underline{v}_1)} = -\delta_{2,1} \quad (7)$$

Das ist aber nichts anderes als der Regressionskoeffizient $\delta_{2,1}$ einer OLS - Regression von $\underline{v}_2$ auf $\underline{v}_1$. Es ist also festzuhalten: OLS ergibt bei Vorliegen einer dreiecksförmigen, d.h. unterhalb der Hauptdiagonalen nur mit Nullen besetzten B -Matrix dann und nur dann konsistente Schätzungen der Koeffizienten dieser B -Matrix und der Koeffizienten der Γ -Matrix, wenn $\beta_{21} = -\delta_{2,1}$ ist. Diese Aussage läßt sich sofort für beliebig große Systeme verallgemeinern, wobei überraschenderweise die Annahme (2) nicht benötigt wird:

Bei einem allgemeinen simultanen Gleichungssystem

$$\underline{Y}_i = -Y_i \beta_i + X \underline{\gamma}_i + \underline{u}_i \quad (i=1, \dots, G) \quad (8)$$

wobei die Matrix Y_i alle endogenen Variablen mit Ausnahme der i -ten enthält und das, mit der Festlegung $\beta_{ii} = 1$, die reduzierte Form besitzt nach

$$Y \cdot I = X \Gamma B^{-1} + U B^{-1} = X \Pi + V \quad (9)$$

können die Koeffizienten der Strukturgleichungen mit OLS konsistent geschätzt werden, wenn gilt

$$\beta_i + \delta_{i,*} = 0 \quad (10)$$

wobei $\delta_{i,*}$ definiert ist nach

$$\delta_{i,*} = (V_i' V_i)^{-1} V_i' \underline{v}_i \quad (11)$$

wobei die Matrix V_i die Matrix der Residuen der reduzierten Form ist, die zu der Matrix Y_i gehört; also

$$V_i = Y_i - X \Pi_i \quad \text{und} \quad (12)$$

$$\underline{v}_i = \underline{y}_i - X \underline{\pi}_i \quad (13)$$

Beweis: Betrachtet man die i -te Strukturgleichung nach (8), so liefert die Anwendung von OLS auf diese Gleichung konsistente Schätzungen für β_i und γ_i , wenn gilt

$$\text{plim}(\frac{1}{T} Y_i' u_i) = 0 \quad (14)$$

Die Matrix

$$((Y_i; X_i)'(Y_i; X_i))^{-1} \text{ existiere,} \quad (15)$$

wobei X_i die Teilmatrix aus X ist, die sich aus den x_j zusammensetzt, für die $\gamma_{ij} \neq 0$ ist.

Y_i läßt sich zerlegen in

$$Y_i = X \Pi_i + V_i \quad (16)$$

Weiterhin sei die übliche Annahme getroffen

$$\text{plim}(\frac{1}{T} X' V) = \text{plim}(\frac{1}{T} X' U) = 0 \quad (17)$$

Somit bleibt die Forderung

$$\begin{aligned} \text{plim}(\frac{1}{T} V_i' u_i) &\stackrel{!}{=} 0 \\ &= \text{plim}(\frac{1}{T} V_i' V \beta_i) \stackrel{!}{=} 0 \end{aligned} \quad (18)$$

Dies aber ist, mit

$$\begin{aligned} \beta_{i+} &= (\beta_{i1}; \beta_{i2}; \dots; \beta_{i, i-1}; \beta_{i, i+1}; \dots; \beta_{i, G}) \quad (19) \\ \Rightarrow \text{plim}(\frac{1}{T} ((V_i' V_i) \beta_{i+} + V_i' v_i)) &\stackrel{!}{=} 0 \\ \Rightarrow \text{plim}(\beta_{i+}) &\stackrel{!}{=} \text{plim}(-(V_i' V_i)^{-1} V_i' v_i) \quad \text{q.e.d.} \quad (20) \end{aligned}$$

Damit läßt sich aber auch v_i darstellen als

$$v_i = V_i (V_i' V_i)^{-1} V_i' v_i + w_i \quad (21)$$

wobei also w_i der Orthogonalanteil von v_i bezüglich dem von der Matrix V_i aufgespannten Raum ist. Und w_i ist ja nichts anderes als u_i .

Diese Beweisführung erscheint dem Verfasser anschaulicher als die von z. B. Johnston zu sein. Außerdem hat sie den noch später zu verwendenden Vorteil, ohne die Annahme (2) auszukommen, die lediglich einen Sonderfall des obigen

Beweises darstellt. Die Ersetzung der sonst zur Definition rekursiver Systeme verwendete Forderung (3), also $\text{cov}(\underline{u}_i, \underline{u}_j) = 0$ für $i \neq j$ in Verbindung mit der Forderung (2) durch die neue Forderung $\text{cov}(\underline{Y}_i, \underline{u}_i) = \text{cov}(\underline{V}_i, \underline{u}_i) = \emptyset$ bedeutet ja nichts anderes als den Übergang von einer hinreichenden, aber nicht nötigen Forderung zu einer notwendigen und hinreichenden Forderung: Der Beweis erlaubt ja ohne weiteres die Existenz von $\text{cov}(\underline{u}_i, \underline{u}_j) \neq 0$ bzw. eine Gestalt der Matrix

B , die von der Dreiecksgestalt abweicht; die gleichzeitige Verletzung der Annahmen (2) und (3) ist allerdings durch Annahme (18) ausgeschlossen.

Vorläufig soll allerdings die Annahme (2) als Modelltypisierung weiterverwendet werden, die erst im nächsten Abschnitt aufgegeben wird. Dies erleichtert zumindest vorläufig die inhaltliche Interpretation der Ergebnisse des obigen Satzes, die sich wie folgt darstellen lassen:

Ein rekursives Kausalketten - System besteht aus einer gerichteten Wirkungsbeziehung, die bei $\underline{x}_1$ beginnt und, nach Durchlaufen des gesamten Systems, bei $\underline{x}_G$ endet. Diese Wirkungsbeziehung findet zwar im ganzen System gleichzeitig statt, eine Analyse wird aber dadurch erleichtert, wenn man sich (beliebig klein wählbare) Zeitabstände zwischen den Realisationen der einzelnen Gleichungen vorstellt. Damit gilt die Sequenz (mit t indizierte Vektoren sind Zeilenvektoren!):

1. Gleichung: $\underline{y}_{1t}$ reagiert auf die Werte von $\underline{x}_{1t}$ zwar im Erwartungswert mit $\underline{x}_{1t} \underline{\gamma}_1$, 'begeht' aber bei der Realisation zum Zeitpunkt t einen Fehler $u_{1t} = v_{1t}$; $E(y_{1t})$ sei η_{1t} .
2. Gleichung: $\underline{y}_{2t}$ reagiert nun auf die Werte von $\underline{x}_{2t}$ im Erwartungswert nach $\underline{x}_{2t} \underline{\gamma}_2$ und auf die Werte von $\underline{y}_{1t}$ im Erwartungswert nach $\underline{y}_{1t} \beta_{21}$. Das heißt aber, daß $\underline{y}_{2t}$ in ihren Reaktionen keinen Unterschied zwischen η_{1t} und v_{1t}

macht: y_{2t} reagiert auf η_{1t} und u_{1t} mit demselben Koeffizienten β_{21} . Damit ist die Größe v_{1t} zu einer Einflußgröße geworden; y_{1t} ist ein stochastischer Regressor geworden. Im Gegensatz zu den in Kapitel 7 behandelten "klassischen" interdependenten Systemen "sieht" y_{2t} also nicht den Erwartungswert von y_{1t} , sondern deren Realisation einschließlich des Fehlers $v_{1t} = u_{1t}$ und reagiert auf diese Realisation.

Betrachtet man schließlich die G-te Gleichung, so gilt: G-te Gleichung: y_{Gt} reagiert auf die Werte von x_{Gt} im Erwartungswert nach $\underline{x}_{Gt} \underline{\gamma}_G$ und auf die Werte von $y_{G,t}$ im Erwartungswert nach $\underline{y}_{G,t} \underline{\beta}_{G+}^1$. Das heißt also, daß $y_{G,t}$ auf eine Matrix stochastischer Regressoren reagiert, deren Elemente y_{it} ($i < G$) bis auf y_{1t} von den Fehlern der jeweils um 1 niedriger indizierten Endogenen abhängig waren und dann, nach dem Erzeugen eines eigenen i-ten Zufallsfehler u_{it} , einschließlich y_{it} diese Überlagerung von 'erhaltenem' und 'erzeugtem' Zufallsfehler der nächsthöher indizierten Variablen y als Regressor 'anboten'.

Damit wird auch der Begriff "Kausalkette" anschaulich, der dieses 'Weiterreichen' von Fehlern niedriger indizierter und der eigenen Strukturgleichung als Regressor beschreibt; der "Grund" für einen Erwartungswert von z. B. $y_{2t} : - \eta_{2t} = y_{1t} \beta_{21} - x_{2t} \underline{\gamma}_2 + \eta_{1t} \beta_{21} - x_{2t} \underline{\gamma}_2$ liegt eben darin, daß y_{2t} nicht zwischen η_{1t} und u_{1t} zu "unterscheiden" vermag. Und solange keine Endogene auf eine niedriger indizierte endogene Variable als Regressor einwirkt, ist diese Kausalkette offen, d. h. eine Rückkopplung auf eine niedriger indizierte Endogene der rekursiven Kette findet nicht statt.

Es ist also deutlich hervorzuheben, daß eine Erfüllung der Bedingung (18) nicht lediglich ein formales Mittel ist, das

1) $\underline{y}_{G,t}$ ist die t-te Zeile von Y_G !

die Identifikation der Strukturgleichungskoeffizienten gewährleistet, sondern daß auch ein gegenüber einem "klassischen" interdependenten System völlig andere inhaltliche Aussagen vorliegen, die auch eine entsprechend unterschiedliche Interpretation verlangen: Dem im interdependenten Fall vorliegenden nur von Realisationen exogener Größen abhängigen Erwartungswert einer Endogenen steht im Fall eines Kausalketten-Modells ein bedingter Erwartungswert gegenüber, der zusätzlich zu den Exogenen auch noch auf eine zufällige Realisation der Fehler in den Gleichungen der als Regressanden auftretenden endogenen Variablen reagiert. Und aufgrund dieser Überlegungen kann man auch zeigen, daß die scheinbar verletzte Abzählregel zur Identifikation einer Strukturgleichung sehr wohl erfüllt ist. Betrachten wir hierzu die i -te Strukturgleichung eines rekursiven Systems, nämlich (mit $i > 1$) :

$$y_{it} = - \sum_{j=1}^{i-1} y_{jt} \beta_{ij} + \sum_{x=1}^K x_{xt} \gamma_{ix} + u_{it} \quad (22)$$

wobei neben den Annahmen (15) und (17) nun die zusätzliche Annahme getroffen werde

$$\beta_{ij} \neq 0 \text{ für alle } j ; \gamma_{ix} \neq 0 \text{ für alle } x \quad (23)$$

d. h. in der i -ten Gleichung ist die maximal mögliche Anzahl von Regressoren. Die Feststellung einer "Nichtidentifizierbarkeit", wie es die Prüfung mit der Abzählregel nach $g-1 > K-k=0$ ergeben würde, ist aber falsch, denn für

$$y_{it} = - \sum_j (\eta_{jt} + v_{jt}) \beta_{ij} + \sum_x x_{xt} \gamma_{ix} + v_{it} + \sum_j v_{jt} \beta_{ij} \quad (24)$$

mit $\eta_{jt} = E(x_t (X'X)^{-1} X' y_j)$ und, wegen (21) , mit

$$v_{it} = - \sum_j v_{jt} \beta_{ij} + w_{it}$$

gilt

$$y_{it} = - \sum_j (\eta_{jt} + v_{jt}) \beta_{ij} + \sum_x x_{xt} \gamma_{ix} + w_{it} \quad (25)$$

Wegen (21) ist $\underline{w}_i$ orthogonal zu den $\underline{v}_j$; da die $\underline{v}_i$ und $\underline{v}_j$ einen zu X orthogonalen Raum aufspannen, ist auch $\underline{w}_i$ als Linearkombination von $\underline{v}_i$ und $\underline{v}_j$ orthogonal zu X . Damit sind die Voraussetzungen für eine konsistente OLS-Schätzung gegeben. Und auch die Abzählregel ist erfüllt: Die Anzahl der aus der i -ten Gleichung ausgeschlossenen Exogenen ist zwar Null, dies trifft jedoch ebenso auf die Anzahl der neben $\underline{y}_i$ auftretenden Endogenen zu: $g - 1 = 0$. Denn die $\underline{y}_j$ sind in der i -ten Gleichung keine endogenen Regressanden mehr, sondern sie sind stochastische Regressoren, die die Bedingungen für konsistente OLS-Schätzungen erfüllen ¹⁾. Die entscheidende Annahme ist hierbei die Bedingung (10), die durch die Umformulierung in (25) anschaulich wird. Wenn y_{it} sowohl auf alle η_{jt} und v_{jt} mit jeweils identischen Regressionskoeffizienten 'reagiert', d. h. auf die von den exogenen Größen induzierten Veränderungen von y_{jt} in gleicher Höhe wie auf die 'autonomen' Fehler von y_{jt} , kann man den Einfluß der beiden Teilgrößen η_{jt} und v_{jt} zu y_{jt} zusammenfassen und mit einer Schätzung identifizieren; die insgesamt $2(g - 1) + K$ Regressoren können wegen der postulierten Gleichheit der Koeffizienten $-\beta_{ij}$ und $\delta_{i,*}$ zu $g - 1 + K$ Regressoren zusammengefaßt werden. Daß diese Gleichheit der Koeffizienten $-\beta_{ij}$ und $\delta_{i,*}$ nicht nur zur Identifizierbarkeit notwendig ist, sondern auch plausibel, soll ein kleines typisches ökonometrisches Beispiel nahelegen. Wenn z. B. gilt

$$C_t = 0.8 Y_t + C_0 + 0.1 C_{t-1} + u_t$$

mit C_t = Konsum einer Periode, Y_t = Einkommen einer Periode,

C_0 = autonomer Konsum, C_{t-1} = Konsum der Vorperiode

und

$$Y_t = \underline{x}_t \underline{\pi} + v_t = \tilde{Y}_t + v_t \quad \text{mit} \quad \tilde{Y}_t = E(Y_t)$$

1) Vgl. Kapitel 1.2.

wobei in $\underline{x}_t$ alle (z. B. vom Staat) gesteuerten exogenen Größen zusammengefaßt sind. Wenn nun C_t der (mikroökonomische) Konsum eines Haushalts ist und durch eine negative Störgröße v_t das Einkommen des Haushaltes gegenüber dem ihm vom Staat zugedachten Einkommen $\tilde{Y}$ erniedrigt wird, so ist doch anzunehmen, daß der Haushalt seinen Konsum an seinem wirklichen Einkommen ausrichtet und nicht an dem, was er 'eigentlich', d. h. im Erwartungswert, an Einkommen beziehen sollte. Ob er zu solch einer Differenzierung zwischen 'geplantem' und 'realisiertem' Einkommen überhaupt in der Lage ist, muß ohnehin bezweifelt werden ¹⁾.

Die Plausibilität dieser Vorstellung des Reaktionsmechanismus' führt ja dann auch dazu, daß bei Vorliegen eines Modells, das die Bedingung (2) erfüllt, in der Praxis immer auch die Annahme (3) getroffen wird. Dies geht sogar soweit, daß Versuche unternommen wurden ²⁾, Modelle mit einer nicht dreiecksförmigen B - Matrix zu einer insgesamt oder doch wenigstens teil -, d. h. blockrekursiven Form überzuführen. Wie aber gezeigt wurde, ist eine derartige Rekursivität für eine konsistente Schätzung mit OLS nicht nötig; allein Annahme (18) ist die notwendige und hinreichende Forderung für eine konsistente Schätzung mit OLS. Diese Annahme ist aber nur für Modelle zu testen, die auch ohne sie identifiziert werden können ³⁾; ein solcher Test wird in Kapitel 9 eingeführt. In allen anderen Fällen reduziert sich eine Prüfung der Annahme (18) auf eine Plausibilitätsprüfung, die aber unbedingt vorgenommen werden sollte. Dabei wird man sinnvollerweise nicht die

1) Das Beispiel ist natürlich eine äußerst simplifizierte ökonomische Betrachtungsweise. Vgl. z. B. Kap. 3.4.1. S. 25
2) Gruber (1977)

3) Für eine Identifikation notwendige Restriktionen können nicht getestet werden! Vgl. z. B. Theil (1971) S. 507
Fußnote 13

Annahme (18) direkt prüfen, sondern die aus ihr folgenden Konsequenz (10), wie dies in der umformulierten Gleichung (25) veranschaulicht wurde. Dies bedeutet für den Ökonometriker die Prüfung, ob er in einer Strukturgleichung den Residuen v_j eine den Erwartungswerten aus der reduzierten Form π_j gleichwertige Rolle als Regressor einräumen kann.

Bei der Anwendung und Interpretation eines solchen Modells ist zu beachten, daß nun zwei verschiedene "Reaktionsformen" des Modells vorliegen: Sowohl Struktur- wie auch reduzierte Form des Systems sind Modelle des Reaktionsverhaltens der Regressanden. Dabei kann die Interpretation der Koeffizienten der reduzierten Form weitgehend von der bei interdependenten Systemen "klassischer" Form übernommen werden. So sagt z. B. die reduzierte Form der zweiten Endogenen aus, daß der Erwartungswert von $y_2 : \pi_2 = X \pi_2$ ist, denn der Erwartungswert von $v_{1t} = u_{1t}$ ist ja nach den üblichen Regressionsannahmen gleich Null. Die Wahl der optimalen Steuerungsinstrumente ist also weiterhin aus der reduzierten Form möglich. Zusätzlich besteht jedoch nun auch eine Steuerungs-, d. h. Eingriffsmöglichkeit über die Strukturform. Denn die Größen v_{jt} können ja auch durch eine bewußt in das System eingebrachte Störung erzeugt werden. Betrachten wir dazu wieder das kleine Beispiel der vorigen Seite und nehmen an, der Staat müsse zur Erreichung bestimmter Ziele die exogenen Größen so wählen, daß über $\tilde{y}_t = x_t \pi$ eine von ihm unerwünschte Einkommenssenkung die Folge wäre, die im Erwartungswert eine ebenso unerwünschte Senkung des Konsums nach sich zöge. Ohne eine Veränderung der Festlegung seiner exogenen Instrumente hat er nun die Möglichkeit, das Einkommen direkt (z. B. durch Steuergeschenke mit einmaligem Charakter) zu beeinflussen und durch das Ein-

bringen einer solchen positiven Störgröße $v_t(Y)$ sowohl Y_t direkt wie auch C_t nach Maßgabe der Strukturgleichung mit dem Faktor 0.8 zu erhöhen. Aus der Strukturform der in dem rekursiven System auf die zur Bestimmung von C_t nachfolgenden Gleichungen (also mit höherem Gleichungsindex) kann er weiter überprüfen, in welcher Form die bedingten Erwartungswerte dieser Endogenen von seinem Eingriff beeinflußt werden, um sich zu veranschaulichen, auf welchem Weg sich sein Eingriff in das System durch die Strukturgleichungen fortpflanzt. Den quantitativen Wert der Auswirkungen seines Eingriffs (bzw. seiner Eingriffe) wird er natürlich an der reduzierten Form überprüfen, denn bezeichnet man eine solche gezielt in die i -te Gleichung eingebrachte Störung mit Δu_{it} , so gilt für die reduzierte Form wegen $V = UB^{-1}$ jetzt auch

$$\Delta V = \Delta UB^{-1} \quad (26.1)$$

bzw., bei Betrachtung eines Zeitpunktes t ¹⁾

$$\Delta \underline{v}_t = \Delta \underline{u}_t B^{-1} \quad (26.2)$$

Betrachtet man z. B. ein solches Modell als eine Beschreibung bestimmter Marktmechanismen, so wird für den eingreifenden Politiker die Auswirkung eines Eingriffs quantitativ überprüfbar, wenn dieser Eingriff nicht mit marktkonformen Mitteln, also mit den exogenen Instrumenten, durchgeführt wird.

8.2.2. Nichtrekursive Systeme

Wie in dem letzten Abschnitt gezeigt wurde, muß ein ökonomisches Gleichungssystem keine rekursive Dreiecksform der B -Matrix besitzen, um eine konsistente Schätzung der Strukturgleichungskoeffizienten mit OLS durchführen zu können.

1) Zur Erinnerung: Mit t indizierte Vektoren sind Zeilenvektoren !

Im numerischen Anhang Seite XII.1 und XII.2 sind zwei kleine exemplarische Fälle aufgeführt, deren Ergebnisse hier für eine Interpretation zumindest auszugsweise dargestellt werden. Schon das erste Modell, das nach der Abzählregel auch schon ohne die Annahme (10), dann allerdings nicht mit OLS, identifiziert werden kann, zeigt wichtige Eigenschaften. Gegeben sei also ein Modell

$$y_1 = 0.5 y_2 - 3.0 x_1 + u_1 \quad (27.1)$$

$$y_2 = 0.5 y_1 - 3.0 x_2 + u_2 \quad (27.2)$$

mit der reduzierten Form

$$y_1 = -4.0 x_1 - 2.0 x_2 + v_1 \quad (28.1)$$

$$y_2 = -2.0 x_1 - 4.0 x_2 + v_2 \quad (28.2)$$

das neben den üblichen Regressionsannahmen (vgl. S. 98) jetzt vor allem auch die Bedingung (10) erfüllt. Die Koeffizienten der reduzierten Form sind wie immer konsistent direkt mit OLS zu schätzen, die Koeffizienten der Strukturgleichung sind aber neben der üblichen Schätzung mit einem konsistenten Schätzer für Strukturgleichungen wie z. B. TSLS wegen (10) auch direkt mit OLS konsistent zu schätzen. Beide Schätzungen stimmen im Wahrscheinlichkeitslimes natürlich überein; das darf aber nicht darüber hinwegtäuschen, daß mit der Wahl der unterschiedlichen Schätzverfahren vollständig verschiedene Modellvorstellungen zugrunde gelegt werden und damit auch völlig unterschiedliche Interpretationen verlangen.

Die Interpretation des Gleichgewichtsmodells im "klassischen" interdependenten Fall wurde in Kapitel 7 hinreichend behandelt und braucht hier nicht wiederholt zu werden; entscheidend sind hier nur die Feststellungen, daß ein solches interdependentes System keine Abhängigkeiten eines endogenen Regressanden von

endogenen Regressoren beschreibt. Das heißt aber insbesondere, daß ein in irgendeiner Gleichung des Systems auftretender Zufallsfehler keine durch die Annahmen des Modells beschriebenen Auswirkungen über diese Gleichung zeigen muß. Denn die "klassischen" Schätzverfahren bestimmen ja lediglich das Verhältnis gleichzeitig auftretender Realisationen von Zufallsgrößen y_i ($i=1, \dots, g_i$) in einer Gleichung neben einem Satz von exogenen Regressanden x_j ($j=1, \dots, k_i$), weisen ihnen aber keine durch die Schätzannahmen begründete Wechselwirkung zu. Anders ist es jedoch, wenn die Bedingung (10) erfüllt ist, dem Modell also eine Kausalkette zugrunde liegt. Dies wird am besten bei der Betrachtung eines Zeitpunktes $t=1$ deutlich, wenn man zu diesem Zeitpunkt eine "Störung" Δv_2 auf die 2. Endogene einwirken läßt. Zur Veranschaulichung und um ein zumindest zeitweiliges Ungleichgewicht zu ermöglichen ¹⁾, sollen die Strukturgleichungen jeweils mit einer bestimmten Zeitverzögerung τ nacheinander als "Reaktionsgleichungen" aufgefaßt werden, also z. B.

$$\begin{aligned} y_{1,t+\tau} &= 0.5 y_{2,t} - 3.0 x_{1,t} + u_{1,t+\tau} \\ y_{2,t+2\tau} &= 0.5 y_{1,t+\tau} - 3.0 x_{2,t} + u_{2,t+2\tau} \\ y_{1,t+3\tau} &= 0.5 y_{2,t+2\tau} - 3.0 x_{1,t} + u_{1,t+3\tau} \quad \text{usw.} \end{aligned}$$

Hier wurden die "Reaktionsverzögerungen" der Einfachheit halber für beide Gleichungen in derselben Höhe angenommen; diese Annahme kann aber ohne Beeinträchtigung der Aussagen fallengelassen werden ²⁾. Die zum Zeitpunkt $t = 1$ auf y_2 einwirkende Störung Δv_2 bei konstantgehaltenen $x_{1,t}$ und $x_{2,t}$ führt dann im Wahrscheinlichkeitslimes wegen $\text{plim}(\sum_{j=1}^T u_i' v_j) = 0$ zu der Sequenz der aufeinanderfolgenden "Reaktionswerte"

1) Vgl. z. B. Neumann (1973) S. 56 ff.

2) Vgl. z. B. Leonhard (1972) S. 77 f.

$y_{12} = y_{11} + 0.5 \Delta v_2 \rightarrow y_{22} = y_{21} + .5^2 \Delta v_2 \rightarrow y_{13} = y_{11} + .5^3 \Delta v_2$
 $\rightarrow y_{23} = y_{21} + .5^4 \Delta v_2$ usw. ¹⁾, das System stabilisiert sich also wieder. Man sieht hier aber sofort, daß durch die Verkopplung der v_i ($i=1, \dots, g_i$) nach (20) bzw. (21) zwangsläufig die Frage nach der Stabilität eines solchen Systems aufgeworfen wird, denn die Existenz der Inversen von B sagt ja nur aus, daß ein Gleichgewicht existieren kann, aber über die Realisierbarkeit, das heißt bei einer Systemstörung über die Konvergenz zu einem solchen Gleichgewicht ist noch keine Aussage getroffen.

Das Vorgehen bei einer solchen Stabilitätsprüfung ist z. B. bei Dhrymes ²⁾ nachzulesen. Eine besonders anschauliche Darstellung der Stabilitätsprüfung liefert Leonhard ³⁾, der neben den numerischen Kriterien z. B. von Hurwitz ⁴⁾ auch graphische Stabilitätsprüfungen anhand der charakteristischen Gleichung von B^+ ⁵⁾ vorführt, die den Einfluß der einzelnen Koeffizienten β_{ij}^+ auf die Stabilität unmittelbar erkennen lassen.

Hat man aufgrund einer solchen Stabilitätsprüfung die Stabilität des gesamten Systems oder zumindest des Teilsystems gefunden, dessen Gleichungen durch die Bedingungen (20) bzw. (21) zu einer geschlossenen ⁶⁾ Kausalkette verknüpft sind, so ist das Modell zumindest nach dem Stabilitätskriterium widerspruchsfrei: Ein solches Modell kann existieren. Zeigt aber eine solche Stabilitätsprüfung Instabilität auf, so kann das Schätzmodell keine vollständige Abbildung der realen Vorgänge sein: Ein in

1) Die aufeinanderfolgenden Zeitpunkte $t + sT$ werden der Einfachheit halber mit einem zweiten Index gekennzeichnet.

2) Dhrymes (1970) S. 521 ff. und auch das Beispiel S. 533 ff.

3) Leonhard (1972) S. 83 ff.

4) Vgl. z. B. Bronstein & Semendjajew (1967) S. 389

5) B^+ ist die Matrix B unter Berücksichtigung der verschiedenen Lags für die β_{ij} im dynamischen Modell.

6) "Geschlossen" durch die Rückkopplungen der Interdependenz, da keine Rekursivität mehr vorliegt.

diesem Modell auftretender Zufallsfehler hätte je nach Größe der Verzögerungszeit τ ¹⁾ mehr oder minder schnell zum "Explo-dieren" des Modells geführt. Damit aber ist das Modell entweder vollständig falsch spezifiziert oder zumindest bezüglich des Regelprozesses unterspezifiziert, der für die Stabilität des wahren Modells sorgt und es damit überhaupt erst existenz-fähig macht.

Damit aber kann die für identifizierbare Modelle testbare Beziehung (20) auch Anlaß für eine Spezifikationsprüfung werden: Wenn eine durch einen Test bestätigte Verkopplung der Zufallsgrößen eines Systems nach (20) vorliegt, so kann man nach den in Abschnitt 8.2 durchgeführten Überlegungen von "gerichteten Kausalbeziehungen" (Regressoren $\rightarrow$ Regressand) ausgehen. Zeigt eine Stabilitätsprüfung anhand der geschätzten Koeffizienten für dieses System Instabilität auf, so kann ein Modell der spezifizierten Form mit den geschätzten Koeffizienten nicht oder, je nach der Größe von τ ¹⁾, nur kurzlebig existieren: In der spezifizierten und geschätzten Form ist eine geschlossene, das heißt durch Interdependenzen rückgekoppelte Kausalkette nicht möglich.

Liegt aber umgekehrt ein die Bedingung (20) erfüllendes, aufgrund der geschätzten Koeffizienten stabiles Modell vor, so ist über diese Bedingung (20) eine Prognose der Auswirkungen einer in eine Gleichung eingebrachte Störung²⁾ möglich, und erwünschte oder unerwünschte Nebenwirkungen in anderen Gleichungen sind abschätzbar: Die Palette der für "optimale Politiken" einsetzbaren Instrumente wird um die als vollwertige Regressoren im Sinne einer Kausalkette auftretenden Endogenen erweitert.

- 1) Dabei muß τ aber immer kleiner als die Meßintervalle sein!
- 2) z. B. dirigistische Eingriffe in das Investitionsverhalten; Eingriffe in das Konsumverhalten ("Rationierungsmarken")

§9. Der Test auf Kausalketten - Beziehungen

Wie das vorangegangene Kapitel gezeigt hatte, kann man bei der Schätzung der Koeffizienten einer Strukturgleichung eines interdependenten Gleichungssystems mit OLS auch dann konsistente Schätzungen erhalten, wenn die B - Matrix des Strukturgleichungssystems nicht rekursiv, das heißt dreiecksförmig ist. Notwendige und hinreichende Bedingung hierfür ist die Orthogonalität der endogenen und exogenen Regressoren der jeweils betrachteten Strukturgleichung zu dem stochastischen Strukturgleichungsresiduum u dieser Gleichung. Weiter war gezeigt worden, daß sich als Folge dieser Bedingung eine Verkopplung der Residuen der reduzierten Form ergibt nach ¹⁾

$$\underline{v}_i = V_i (V_i' V_i)^{-1} V_i' \underline{v}_i + \underline{w}_i \quad (1)$$

bzw. wegen der Gleichheit

$$\text{plim } (\underline{\beta}_{i+}) = - \text{plim} ((V_i' V_i)^{-1} V_i' \underline{v}_i) \quad (2)$$

die Beziehung

$$\underline{v}_i = - V_i \underline{\beta}_{i+} + \underline{w}_i = - V_i \underline{\beta}_{i+} + \underline{u}_i \quad (3)$$

Durch diese Gleichheit (2) konnte man den endogenen Regressoren innerhalb der Gleichungen die Rolle von beeinflussenden Größen, dem endogenen Regressand die Rolle einer beeinflussten Größe zuordnen. Damit beschreibt aber eine Strukturgleichung ein Reaktionsverhalten einer endogenen Variablen auf Zustände oder Zustandsänderungen der zugehörigen endogenen und exogenen Regressoren. In Anlehnung an das angelsächsische Schrifttum wird dann von einer "Kausalbeziehung" gesprochen und bei einer Verknüpfung mehrerer solcher "Kausalbeziehungen" von einer "Kausalkette" ("causal chain"). In Umkehrung der Argumentation

1) Vgl. Seite 127 Gleichungen (20) und (21)

aber kann man dann von einer "Kausalbeziehung" sprechen, wenn die Bedingung (2) erfüllt ist. Das Vorliegen einer solchen "Kausalbeziehung" hat für den Anwender der Schätzergebnisse unmittelbare Bedeutung, da dann die Auswirkung von störenden Eingriffen an den endogenen Regressanden prognostizierbar wird. Zum Test auf eine solche "Kausalbeziehung" dient der folgende F - Test, der die Beziehung (2) als Null-Hypothese prüft.

Wegen der unterschiedlichen Bedeutung von endogenen Regressoren und endogenem Regressand wird eine von der üblichen abweichende Nomenklatur überschaubarer, die sich an die Nomenklatur des letzten Kapitels anlehnt.

Beobachtet werde o. B. d. A. die erste Strukturgleichung

$$\underline{Y}_1 = Y_1 \underline{\beta}_1 + X_1 \underline{\gamma}_1 + \underline{u}_1 \quad (4)$$

$$T \begin{bmatrix} 1 \\ \underline{Y}_1 \end{bmatrix} = T \begin{bmatrix} g-1 \\ Y_1 \end{bmatrix} \begin{bmatrix} 1 \\ \underline{\beta}_1 \end{bmatrix} + T \begin{bmatrix} k \\ X_1 \end{bmatrix} \begin{bmatrix} 1 \\ \underline{\gamma}_1 \end{bmatrix} + T \begin{bmatrix} 1 \\ \underline{u}_1 \end{bmatrix} \quad \begin{matrix} (g \hat{=} g_1) \\ (k \hat{=} k_1) \end{matrix}$$

wobei $\underline{Y}_1$ der $T \cdot 1$ - Vektor des endogenen Regressanden

Y_1 die $T \cdot (g_1 - 1)$ -Matrix der endogenen Regressoren ¹⁾

X_1 die $T \cdot k_1$ - Matrix der exogenen Regressoren

$\underline{\beta}_1$ der $(g_1 - 1) \cdot 1$ - Vektor der Koeffizienten der endogenen Regressanden

$\underline{\gamma}_1$ der $k_1 \cdot 1$ - Vektor der Koeffizienten der exogenen Regressanden

und $\underline{u}_1$ der stochastische Störterm der ersten Gleichung sind.

Weiter werde angenommen, daß die u_{1t} ($t=1, \dots, T$) identisch

$(0, \sigma^2)$ -normalverteilt sind und orthogonal zu allen Elementen

1) Hierbei ist also der Vektor $\underline{\beta}_1$ um β_{11} gekürzt und hat

deshalb nur noch die Länge $g_1 - 1$! Ebenso ist das Vorzeichen

gegenüber der Darstellung auf S. 124 zur Vereinfachung geändert.

aus X_1 und aus X , wobei X die Matrix der Exogenen des gesamten Gleichungssystems ist.

Der zu der Gleichung (4) zugehörige Teil der reduzierten Form des vollständigen Systems lautet dann

$$\underline{Y}_1 = X \underline{\pi}_1 + \underline{v}_1 \quad (5)$$

und

$$\begin{array}{c} Y_1 = X \Pi_1 + V_1 \\ \begin{array}{|c|c|} \hline 1 & g-1 \\ \hline \end{array} \begin{array}{|c|c|} \hline T & T \\ \hline \end{array} \begin{array}{|c|c|} \hline \underline{Y}_1 & Y_1 \\ \hline \end{array} = \begin{array}{|c|c|} \hline T & X \\ \hline \end{array} \begin{array}{|c|c|} \hline 1 & g-1 \\ \hline \end{array} \begin{array}{|c|c|} \hline \underline{\pi}_1 & \Pi_1 \\ \hline \end{array} K \\ \begin{array}{|c|c|} \hline 1 & g-1 \\ \hline \end{array} \begin{array}{|c|c|} \hline T & T \\ \hline \end{array} \begin{array}{|c|c|} \hline \underline{v}_1 & V_1 \\ \hline \end{array} \end{array} \quad (6)$$

wobei X die $T \cdot K$ - Matrix aller im System vorhandenen Exogenen

$\underline{\pi}_1$ der $K \cdot 1$ - Vektor der Koeffizienten der reduzierten Form der 1. Gleichung

Π_1 die $K \cdot (g-1)$ - Matrix der Koeffizienten der reduz. Form der endogenen Regressoren der 1. Gl.

$\underline{v}_1$ der $T \cdot 1$ - Vektor des stochastischen Residuums der reduz. Form der 1. Gleichung

V_1 die $T \cdot (g-1)$ - Matrix der stochastischen Residuen der reduz. Form der endogenen Regressoren der 1. Gleichung

ist. Identitäten seien in dem vollständigen System nicht enthalten. Die Matrix X läßt sich partitionieren nach

$$X = (X_1, X_2) = \begin{array}{c} \begin{array}{|c|c|} \hline k_1 & (K-k_1) \\ \hline \end{array} \\ \begin{array}{|c|c|} \hline T & T \\ \hline \end{array} \begin{array}{|c|c|} \hline X_1 & X_2 \\ \hline \end{array} \end{array} \quad (7)$$

wobei X_2 die $T \cdot (K - k_1) = T \cdot k_2$ - Matrix der aus der ersten Gleichung ausgeschlossenen Exogenen umfaßt.

Weiter seien die v_{it} identisch für alle $t = 1, \dots, T$

$(0, \sigma_{v_i}^2)$ -normalverteilt für $i=1, \dots, g$. Damit erhält man

die Varianz-Kovarianz - Matrix der Residuen der reduzierten

Form der an der ersten Strukturgleichung beteiligten Endogenen

$$\Sigma_v = \Omega = \begin{bmatrix} \omega_{11} & | & \omega_{12} \\ \hline \omega_{21} & | & \Omega_{22} \end{bmatrix} \quad (8)$$

wobei diese Matrix in Analogie zu (6) ebenfalls partitioniert wurde. Für den Test sind noch nötig die Umformungen

$$\underline{u}_1 = \underline{v}_1 - V_1 \underline{\beta}_1 \quad (9)$$

und eine "Hilfsregression" nach

$$\underline{v}_1 - \underline{w} = V_1 \underline{b}_1 = V_1 (V_1' V_1)^{-1} V_1' \underline{v}_1 = V_1 \Omega_{22}^{-1} \omega_{21} \quad (10)$$

d. h. der $(\mathbb{E}_1 - 1) \cdot 1$ - Vektor $\underline{b}_1$ ist eine OLS-Schätzung, sodaß der Ausdruck $V_1 \underline{b}_1$ den durch V_1 erklärten Anteil von $\underline{v}_1$ darstellt, also den durch V_1 aufgespannten Raum für $\underline{v}_1$, während $\underline{w}$ den zu V_1 orthogonalen ⁽³⁾ unerklärten Anteil von $\underline{v}_1$ darstellt. Dann ist auch w_t ($0, \sigma_w^2$) identisch normalverteilt für alle $t=1, \dots, T$ mit

$$\sigma_w^2 = \omega_{11} - \omega_{12} \Omega_{22}^{-1} \omega_{21} \quad (11)$$

Mit (9) und (10) erhält man

$$\underline{u}_1 = V_1 \underline{b}_1 + \underline{w} - V_1 \underline{\beta}_1 = V_1 (\underline{b}_1 - \underline{\beta}_1) + \underline{w} \quad (12)$$

d. h. $\underline{u}_1$ läßt sich orthogonal zerlegen in den von V_1 erklärten Anteil und in den zu V_1 orthogonalen Anteil von $\underline{u}_1$, nämlich $\underline{w}_1$. Einsetzen von (12) in (4) ergibt daher

$$\underline{y}_1 = Y_1 \underline{\beta}_1 + X_1 \underline{\gamma}_1 + V_1 (\underline{b}_1 - \underline{\beta}_1) + \underline{w} \quad (13)$$

Damit sieht man, daß $\underline{u}_1$ nur dann mit V_1 und damit mit Y_1 unkorreliert ist, wenn $\underline{b}_1 = \underline{\beta}_1$ ist. Ist diese Bedingung jedoch erfüllt, so kann die betrachtete Strukturgleichung mit OLS konsistent geschätzt werden. Die Testhypothese ist also

- (1) Das Vorzeichen bei $\underline{\beta}_1$ ist Folge der Darstellungsform (4)!
- (2) Für den vollen Rang von Ω_{22} war der Ausschluß der Identitäten nötig.
- (3) Als Folge der Normalengleichungen
- (4) Diese Varianzzerlegung ist wegen der Orthogonalität von V_1 und $\underline{w}$ möglich.

$$H_0 : \underline{b}_1 = \underline{\beta}_1 \quad (14)$$

Ersetzt man nun V_1 in (13) durch die Teilmatrix aus (6), so erhält man

$$\underline{Y}_1 = Y_1 \underline{\beta}_1 + X_1 \underline{\gamma}_1 + (Y_1 - X_1 \Pi_1)(\underline{b}_1 - \underline{\beta}_1) + \underline{w} \quad (15)$$

bzw. nach der in (7) durchgeführten Aufspaltung von X in X_1 und X_2 und entsprechender Aufspaltung von Π_1 nach

$$\Pi_1 = (\Pi_{11} ; \Pi_{12})' = \begin{pmatrix} \Pi_{11} \\ \Pi_{12} \end{pmatrix} \quad (16)$$

kommt man zu

$$\begin{aligned} \underline{Y}_1 &= Y_1 \underline{\beta}_1 + X_1 \underline{\gamma}_1 + Y_1 \underline{b}_1 - Y_1 \underline{\beta}_1 - X_1 \Pi_{11} \underline{b}_1 - X_2 \Pi_{12} \underline{b}_1 \\ &\quad + X_1 \Pi_{11} \underline{\beta}_1 + X_2 \Pi_{12} \underline{\beta}_1 + \underline{w} \end{aligned} \quad (17)$$

$$= Y_1 \underline{b}_1 + X_1 (\underline{\gamma}_1 - \Pi_{11} (\underline{b}_1 - \underline{\beta}_1)) - X_2 \Pi_{12} (\underline{b}_1 - \underline{\beta}_1) + \underline{w} \quad (17)$$

Diese Gleichung (17) ist Ausgangsgleichung für das Testen von H_0 . Denn wegen der Rangregel $\text{rg}(\Pi_{12}) = g_1 - 1$ muß Π_{12} den vollen Rang besitzen¹. Damit verschwindet der Vektor $\Pi_{12}(\underline{b}_1 - \underline{\beta}_1)$ nur dann, wenn $(\underline{b}_1 - \underline{\beta}_1) = \emptyset$ ist, wenn also H_0 zutrifft. Da (17) eine lineare Gleichung mit stochastischen Regressoren ist, deren Fehlerterm $\underline{w}$ mit allen Regressoren unkorreliert ist² und damit auch jede OLS - Schätzvariable dieser Gleichung stochastisch unabhängig von jedem $\hat{\underline{w}}$ ist.³, kann man den bekannten Joint - Test anwenden nach

$$\frac{(\hat{\underline{w}}' \hat{\underline{w}} - \hat{\underline{w}}' \hat{\underline{w}}) / (K - k_1)}{(\hat{\underline{w}}' \hat{\underline{w}}) / (T - K - g_1 + 1)} = \frac{(\hat{\underline{Y}} - \hat{\underline{Y}})' (\hat{\underline{Y}} - \hat{\underline{Y}}) / (K - k_1)}{(\hat{\underline{w}}' \hat{\underline{w}}) / (T - K - g_1 + 1)} \quad (18)$$

ist F - verteilt mit $(K - k_1), (T - K - g_1 + 1)$ Freiheitsgraden, wobei die verwendeten Kurzbezeichnungen bedeuten :

- 1) Hier wird deutlich, daß nur identifizierbare Gleichungen getestet werden können! Vgl. auch Seite 112 Fußnote 1
- 2) Und wegen der Normalverteilung auch unabhängig
- 3) Vgl. z. B. Schneeweiß (1974) Seite 112 f.

$\hat{\underline{y}}$ ist die mit OLS geschätzte Gerade für $\underline{y}$ unter Verwendung aller $g-1$ endogenen und K exogenen Regressoren Y_1, X_1, X_2

$\hat{\underline{w}}$ ist das zugehörige OLS-Residuum nach $\hat{\underline{w}} = \underline{y} - \hat{\underline{y}}$

$\hat{\underline{y}}$ ist die mit OLS geschätzte Gerade für $\underline{y}$ unter Verwendung lediglich der $g-1$ endogenen und k_1 exogenen Regressoren Y_1 und X_1 , also unter Ausschluß von X_2

$\hat{\underline{w}}$ ist das zugehörige OLS-Residuum nach $\hat{\underline{w}} = \underline{y} - \hat{\underline{y}}$.

Die Herleitung dieses Joint - Tests findet man für den Fall der multiplen Eingleichungsregression bei Schneeweiß ¹⁾; dieser führt vor allem die sehr anschauliche Form (18) des Joint - Tests ein. Dort findet man auch den Hinweis auf die Machtfunktion dieses Tests, die auf einer nichtzentralen F-Verteilung beruht und z. B. bei Schönfeld ²⁾ nachzulesen ist; dieser führt auch den Beweis der Zulässigkeit der Verwendung des F-Tests bei stochastischen Regressoren ³⁾.

Der hier eingeführte Test scheint dem Verfasser anschaulicher und vor allem rechentechnisch einfacher zu sein als die von De-Min Wu ⁴⁾ vorgelegten Tests auf Unabhängigkeit zwischen dem Störglied $\underline{u}$ und den stochastischen Regressoren Y_1 . Auch die Tests von Wu sind F-Tests, die aber im Gegensatz zu obigem Test nicht das Verschwinden der zu X_2 gehörigen OLS-Koeffizienten testen, sondern unmittelbar die Gleichheit von $\hat{\beta}_{IV}$ als Instrumentalvariablenschätzer (z. B. mit 2SLS) für β_1 und $\hat{\beta}_{OLS}$ als OLS-Schätzer für β_1 testen. Dadurch wird der Rechenaufwand bei Wu nicht unerheblich größer, denn im Zähler aller seiner vier alternativen Testgrößen steht ein Ausdruck

1) Schneeweiß (1974) S. 112 ff., besonders aber S. 116

2) Schönfeld (1969) Band I S. 116 f.

3) Schönfeld (1971) Band II S. 22 f.

4) De-Min Wu (1973)

$$Q = (\hat{\beta}_{OLS} - \hat{\beta}_{IV})' ((Y_1' A_2 Y_1)^{-1} - (Y_1' A_1 Y_1)^{-1})^{-1} (\hat{\beta}_{OLS} - \hat{\beta}_{IV})$$

mit $A_1 = I - X_1(X_1' X_1)^{-1} X_1'$ und

$$A_2 = X(X' X)^{-1} X' - X_1(X_1' X_1)^{-1} X_1' .$$

Dieses häufige Invertieren stellt bei den in der Ökonometrie ja oft hochkorrelierten Variablen und damit fastsingulären Momentmatrizen nach Meinung des Verfassers eine nicht unerhebliche Fehlerquelle dar. Darüber hinaus äußert sich Wu nicht zu dem Ziel eines solchen Tests; ohne eine solche Zielsetzung stellt sich sein Ansatz lediglich als Test dar, der nach der Durchführung z. B. einer 2SLS-Schätzung unter Verwendung der daraus erhaltenen Schätzkoeffizienten nachträglich prüft, ob man nicht auch mit OLS hätte konsistent schätzen können. Dagegen prüft der vorstehend eingeführte Test des Verfassers zwar inhaltlich übereinstimmend die Orthogonalität der endogenen Regressoren bezüglich der Strukturgleichungsresiduen, er führt diese Prüfung aber unmittelbar in einer OLS-Schätzung durch und gibt vor allem der Nullhypothese über die formale Definition hinaus einen interpretatorischen Inhalt.

Bei den Ausdrücken aller mit OLS^+ berechneten Beispiele des Anhangs bezieht sich der Ausdruck "Joint-Test auf Unabhängigkeit mit ... Freiheitsgraden : $F^+ = \dots$ " auf obigen Joint-Test. Eine Aussage aber hat er lediglich bei den Beispielen des Anhangs A.XII, bei denen Modelle mit stochastischen Regressoren gegeben sind im Sinne "kausaler" Beziehungen. Daß bei den Beispielen F^+ den Wert Null annimmt, ist ja durch die Konstruktion der Modelle bedingt, die die Bedingung (2) erfüllen.

§ 10. Schlußbemerkungen

=====

In Kapitel 6 waren die zur konsistenten Regressionsschätzung des 'EIV'-Modells und des 'unterspezifizierten' Modells notwendigen zusätzlichen Informationen aus Gleichungen eines Gleichungssystems gewonnen worden, die bis auf eine gemeinsame Fehlspezifikation ¹⁾ keine formale Interdependenz aufwiesen ('SUR-Systeme'). Mit dem Ziel, weitere Zusatzinformationen zu erhalten, wurden deshalb anschließend die Modellvorstellungen interdependenter Systeme untersucht. Dabei stellte sich heraus, daß die Strukturgleichungen interdependenter Systeme "klassischer" Prägung weder wechselseitige kausale Wirkungsbeziehungen zwischen den Gleichungen noch zwischen den endogenen Variablen einer Strukturgleichung abbilden, sondern lediglich das Verhältnis gleichzeitig gemessener Realisationen von Zufallsgrößen y_i entsprechend der komparativ-statischen Modellvorstellung des Ökonometrikers quantifizieren: Die zu den Endogenen gehörenden Strukturgleichungskoeffizienten sind partielle Korrelationskoeffizienten zwischen Zufallsvariablen, aber keine Regressionskoeffizienten, die eine gerichtete Beziehung zwischen Regressand und Regressoren abbilden. Damit aber stellen die Strukturgleichungsresiduen u_i keine "Verhaltensfehler" als Folge gestörter Reaktionen von y_i auf die $y_{i'}$ ($i' = 1, \dots, g_i; i \neq i'$) und die x_j ($j = 1, \dots, k_i$) dar, sondern sind lediglich die Abbildung des nicht durch den Raum aller Exogenen X des Systems erklärten stochastischen Prozesses, wie er sich in y_i nach Maßgabe von β_i überlagert. Damit jedoch kann der Verfasser der Argumentation von Schneeweiß ²⁾ nicht folgen, der bei einer vollständig spezifizierten Struktur-

1) Also Spezifikation mit Proxies bzw. Unterspezifikation

2) Schneeweiß (1974) S. 253

gleichung von einer 'mutual correlation' zwischen den Strukturgleichungen $r(\underline{u}_i, \underline{u}_{i'}) = 0$ ($i, i' = 1, \dots, G; i \neq i'$) ausgeht und eine nichtverschwindende Korrelation als Indiz für Fehlspezifikation ansieht. Da aber eine verschwindende Kovarianz zwischen den Gleichungsresiduen für die Schätzansätze des Kapitels 6 notwendige Bedingung war, ist für eine Übertragung dieser Ansätze auf interdependente Systeme nach der Überzeugung des Verfassers keine modelltheoretische Begründung möglich. Das gleiche gilt für die nichtrekursiven Modelle des Kapitels 8. Lediglich die rekursiven Modelle nach Abschnitt 8.2.1. erfüllen durch ihre Modellannahmen die Forderung der verschwindenden Kovarianz zwischen den Residuen der einzelnen Strukturgleichungen. Damit lassen sich zwar die Ansätze des Kapitels 6 unverändert auf rekursive Systeme übertragen, da die Einzelgleichungen nicht nur "seemingly unrelated", sondern durch die Modellannahmen "unrelated" sind. Zusätzliche Informationen aber liefern auch diese Systeme durch die Einbeziehung von endogenen Variablen in eine Gleichung als stochastische Regressoren nicht.

Den Abschluß sollen noch einige Bemerkungen zu den dargestellten Eigenschaften der in dieser Arbeit vorgeführten Schätzansätze bilden. Für einen Statistiker, der von den "BLUE"-Eigenschaften des OLS-Schätzers ausgeht, mag es enttäuschend sein, wenn so gut wie alle Schätzansätze lediglich konsistent, aber nicht unverzerrt formuliert wurden. Denn die daraus ableitbaren, nur asymptotischen Eigenschaften sind nach Johnston ¹⁾ "cold comfort" für einen Ökonometriker, der ja mit endlichen Stichproben arbeiten muß. Diese Eigenschaften sind jedoch Abbild des unzulänglichen Datenmaterials, mit dem bei

1) Johnston (1963) S. 275

empirischen Untersuchungen im wirtschafts- und sozialwissenschaftlichen Bereich gearbeitet werden muß. Und bei dem für dieses Anwendungsgebiet typischen Datenmaterial kann auch für OLS die "BLUE"-Eigenschaft nicht mehr aufrecht erhalten werden ¹⁾. Der Verfasser hält es auf jeden Fall für sinnvoller, wenn durch die Verwendung von "schwachen" Annahmen entsprechend schwache Eigenschaften eines Schätzers formuliert werden und damit auch zu entsprechender Vorsicht bei der Interpretation der Schätzergebnisse anhalten, als wenn man sich durch nicht den Daten oder dem Modell adäquate Annahmen dem Selbstbetrug "bester" Eigenschaften des Schätzers hingibt: Mit einer nur definierten, aber nicht problemkonformen "Optimalität" eines Schätzers ist niemandem gedient.

1) Vgl. z. B. den Abschnitt 1.2. "Stochastische Regressoren" .

A N H A N G

= = = = =

(Numerische Beispiele)

Von OLS* verwendete Daten fuer :

```

y : -3.0000 -11.0000 12.0000 3.0000 8.0000 -9.0000
x1 : -2.0000 -1.0000 0.0000 0.0000 1.0000 2.0000
p1 : -1.0000 -2.0000 1.0000 -1.0000 0.0000 3.0000
x2 : 0.0000 -1.0000 1.0000 1.0000 0.0000 -1.0000
x3 : 0.0000 -1.0000 2.0000 0.0000 1.0000 -2.0000

```

$$p_{1t} = x_{1t} + v_t^+ = x_{1t}^+$$

```

y = 2.0000 x1 3.0000 x2 5.0000 x3
( 5.3666) ( 3.1820) ( 8.3205)

```

s**2 : 1.333 () d : 1.750 () B(adj.) : .9844 () Fr.-grade : 3 () "t"-Werte : Werte in Klammern

Joint-Test auf Unabhaengigkeit mit 1, 2 Freiheitsgraden : F* : 0.00

Prognosen aus Strukturform fuer :

```

y : -4.0000 -10.0000 13.0000 3.0000 7.0000 -9.0000

```

Prognosefehler aus Strukturform (=U) fuer :

```

y : 1.0000 -1.0000 -1.0000 0.0000 1.0000 0.0000 = v_t

```

```

y = 1.2308 p1 3.0000 x2 4.8462 x3
( 1.9435) ( 1.4669) ( 3.7340)

```

s**2 : 6.256 () d : 1.399 () B(adj.) : .9269 () Fr.-grade : 3 () "t"-Werte : Werte in Klammern

Joint-Test auf Unabhaengigkeit mit 1, 2 Freiheitsgraden : F* : 7.38

```

y = 3.0000 x2 4.6000 x3
( 1.1285) ( 2.7360)

```

s**2 : 10.600 () d : .520 () B(adj.) : .8762 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern

Joint-Test auf Unabhaengigkeit mit 2, 2 Freiheitsgraden : F* : 9.60

```

y = 2.0000 x1 .0000 p1 3.0000 x2 5.0000 x3
( 2.7175) ( .0000) ( 2.5981) ( 6.7937)

```

s**2 : 2.000 () d : 1.750 () B(adj.) : .9766 () Fr.-grade : 2 () "t"-Werte : Werte in Klammern

Von OLS* verwendete Daten fuer :

```
y : -7.0000  0.0000 -9.0000 -1.0000 -2.0000 19.0000
x1 : -2.0000 -1.0000  0.0000  0.0000  1.0000  2.0000
x2 : -2.0000  0.0000 -2.0000  0.0000  0.0000  4.0000
p2 : -2.0000 -1.0000 -2.0000  2.0000 -1.0000  4.0000
```

```
y = -1.0000 x1  5.0000 x2
    (-1.6330) (12.6491)
```

```
s**2 : 1.500 () d : 2.667 () B(adj.) : .9649 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 1, 3 Freiheitsgraden : F* : .00
```

Prognosen aus Strukturform fuer :

```
y : -8.0000  1.0000 -10.0000  0.0000 -1.0000 18.0000
```

Prognosefehler aus Strukturform (=U) fuer :

```
y : 1.0000 -1.0000  1.0000 -1.0000 -1.0000  1.0000
```

```
y = 1.3077 x1  3.0769 p2
    (.6015) ( 2.4514)
```

⊙)

```
s**2 : 24.577 () d : 1.466 () B(adj.) : .7522 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 1, 3 Freiheitsgraden : F* : 46.16
```

```
y = 5.0000 x1
    ( 2.2542)
```

⊙)

```
s**2 : 49.200 () d : 2.250 () B(adj.) : .5040 () Fr.-grade : 5 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 2, 3 Freiheitsgraden : F* : 60.00
```

- ⊕) Man beachte vor allem, daß bei der Spezifikation lediglich mit x_1 der Koeffizient der OLS-Schätzung schon mit 90% Sicherheitswahrscheinlichkeit signifikant positiv von Null verschieden ist ($t_{f=5}^{90\%} = 2.015$), während dies bei dem einschließlich der Proxy spezifizierten Ansatz wegen der großen Varianz von $\hat{\gamma}_1$ nicht der Fall ist !

Von OLS* verwendete Daten fuer :

Y : -9.0000 -6.0000 1.0000 -1.0000 4.0000 11.0000 $\hat{=}$ unregelter Prozess
 x1 : -2.0000 -1.0000 0.0000 0.0000 1.0000 2.0000

Y = 5.0000 x1
 (14.4338)

s**2 : 1.200 () d : 2.567 () B(adj.) : .9766 () Fr.-grade : 5 () "t"-Werte : Werte in Klammern

Prognosen aus Strukturform fuer :

Y : -10.0000 -5.0000 0.0000 0.0000 5.0000 10.0000 $\hat{=}$ Zielgröße der Regelung

Prognosefehler aus Strukturform (=U) fuer :

Y : 1.0000 -1.0000 1.0000 -1.0000 -1.0000 1.0000

Von OLS* verwendete Daten fuer :

eta : -10.0000 -5.0000 0.0000 0.0000 5.0000 10.0000 $\hat{=}$ geregelter Prozess
 x+K : -2.2000 -.8000 -.2000 .2000 1.2000 1.8000

eta = 4.8828 x+K
 (14.4338)

s**2 : 1.172 () d : 2.592 () B(adj.) : .9766 () Fr.-grade : 5 () "t"-Werte : Werte in Klammern

Prognosen aus Strukturform fuer :

eta : -10.7422 -3.9062 -.9766 .9766 5.8594 8.7891

Prognosefehler aus Strukturform (=U) fuer :

eta : .7422 -1.0937 .9766 -.9766 -.8594 1.2109

Von SURE verwendete Daten fuer :

```

y1 : -7.0000 -5.0000 -1.0000 1.0000 1.0000 11.0000
y2 : 14.0000 6.0000 2.0000 -7.0000 -2.0000 -13.0000

x1 : -2.0000 -1.0000 0.0000 0.0000 1.0000 2.0000
x2 : 2.0000 2.0000 -1.0000 -1.0000 -1.0000 -1.0000
x3 : -4.0000 -2.0000 -2.0000 2.0000 0.0000 6.0000

```

Gamma OLS fuer :

```

y1 = 2.0000 x1 + 1.0000 x3
      (.6405) (.2532)

```

```

s**2 : 1.000 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern
-----

```

```

y2 = 2.0000 x2 -2.0000 x3
      (.4901) (.2122)

```

```

s**2 : 1.667 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern
-----

```

R = Korrelationsmatrix der Residuen :

```

1.0000 .7746
.7746 1.0000

```

Inverse von R :

```

2.5000 -1.9365
-1.9365 2.5000

```

Gamma des SUR-Schaetzers nach 0 Iterationen fuer:

```

y1 = 2.0000 x1 + 1.0000 x3
      (.4776) (.2064)

```

```

s**2 : 1.000 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern
-----

```

```

y2 = 2.0000 x2 -2.0000 x3
      (.3655) (.1913)

```

```

s**2 : 1.667 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern
-----

```

Residuen des SUR-Schaetzers fuer:

```

y1 : 1.0000 -1.0000 1.0000 -1.0000 -1.0000 1.0000
y2 : 2.0000 -2.0000 -.0000 -1.0000 -.0000 1.0000

```

R der SURE-Residuen :

```

1.0000 .7746
.7746 1.0000

```

Maximale rel. Abweichung : .2861022E-06 nach 0 Iterationen

Gamma OLS fuer :

y1 = 4.2000 x1
(.6000)

s**2 : 3.600 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

y2 = 5.0000 x2
(1.4814)

s**2 : 26.333 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

R = Korrelationsmatrix der Residuen :

1.0000 -.5238
-.5238 1.0000

Inverse von R :

1.3781 .7219
.7219 1.3781

Gamma des SUR-Schaetzers nach 0 Iterationen fuer:

y1 = 3.8414 x1
(.5662)

s**2 : 3.814 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

y2 = 4.8315 x2
(1.3980)

s**2 : 26.390 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

y1 : .6828 -1.1586 -1.0000 1.0000 -2.8414 3.3172
y2 : 4.3371 -3.6629 6.8315 -2.1685 2.8315 -8.1685

R der SURE-Residuen :

1.0000 -.6136
-.6136 1.0000

Maximale rel. Abweichung : .2070935E+00 nach 0 Iterationen

Gamma des SUR-Schaetzers nach 12 Iterationen fuer:

y1 = 3.6727 x1
(.5721)

s**2 : 4.063 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

y2 = 4.5895 x2
(1.3781)

s**2 : 26.670 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

y1 : .3654 -1.3273 -1.0000 1.0000 -2.6727 3.6546
y2 : 4.8210 -3.1790 6.5895 -2.4105 2.5895 -8.4105

R der SURE-Residuen :

1.0000 -.6528
-.6528 1.0000

Maximale rel. Abweichung : .3157624E-05 nach 12 Iterationen

Von OLS* verwendete Daten fuer :

```

y1 : -3.0000 -8.0000 11.0000 -1.0000 6.0000 -5.0000
y2 : -5.0000 -14.0000 19.0000 0.0000 14.0000 -14.0000

x1 : -2.0000 -1.0000 0.0000 0.0000 1.0000 2.0000
x2 : 0.0000 -1.0000 2.0000 0.0000 1.0000 -2.0000
p2 : 0.0000 -2.0000 2.0000 2.0000 0.0000 -2.0000

```

```

-----
y1 = 2.0000 x1 5.0000 x2 (1)
      ( 5.0595) ( 12.5491)

```

```

s**2 : 1.500 () d : 2.667 () B(adj.) : .9707 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 1, 3 Freiheitsgraden : F* : .00
-----

```

```

y2 = 3.0000 x1 10.0000 x2 (2)
      ( 9.2952) ( 30.9839)

```

```

s**2 : 1.000 () d : 1.750 () B(adj.) : .9949 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 1, 3 Freiheitsgraden : F* : .00
-----

```

Korrelationsmatrix der U :

```

1.0000 0.0000
0.0000 1.0000

```

Prognosen aus Strukturform fuer :

```

y1 : -4.0000 -7.0000 10.0000 0.0000 7.0000 -6.0000
y2 : -6.0000 -13.0000 20.0000 0.0000 13.0000 -14.0000

```

Prognosefehler aus Strukturform (=U) fuer :

```

y1 : 1.0000 -1.0000 1.0000 -1.0000 -1.0000 1.0000
y2 : 1.0000 -1.0000 -1.0000 0.0000 1.0000 -1.0000

```

```

-----
y1 = 1.6154 x1 3.0769 p2 (3)
      ( 1.0175) ( 2.4514)

```

```

s**2 : 24.577 () d : 3.500 () B(adj.) : .5200 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 1, 3 Freiheitsgraden : F* : 45.15
-----

```

```

y2 = 2.2308 x1 6.1538 p2 (4)
      ( .7211) ( 2.5162)

```

```

s**2 : 93.308 () d : 3.486 () B(adj.) : .5210 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 1, 3 Freiheitsgraden : F* : 276.95
-----

```

Korrelationsmatrix der U :

```

1.0000 .9638
.9638 1.0000

```

```
-----
y1 = 1.0000 x1 (5)
      (.4508)

s**2 : 49.200 () d : 2.760 () B(adj.) : .0391 () Fr.-grade : 5 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabh angigkeit mit 2, 3 Freiheitsgraden : F* : 60.00
-----
```

```
-----
y2 = 1.0000 x1 (6)
      (.2277)

s**2 : 192.000 () d : 2.588 () B(adj.) : .0103 () Fr.-grade : 5 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabh angigkeit mit 2, 3 Freiheitsgraden : F* : 350.05
-----
```

Korrelationsmatrix der U :

```
1.0000 .9857
.9857 1.0000
```

Fuer p2 sind Instrumentalvariable: y2

```
y1 = 2.0000 x1 5.0000 p2 (7)
      ( 5.0592) (12.6278)

s**2 : 1.5000 () B(adj.) : .970703 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern
-----
```

Fuer p2 sind Instrumentalvariable: y1

```
y2 = 3.0000 x1 10.0000 p2 (8)
      ( 9.2907) (30.5042)

s**2 : 1.0000 () B(adj.) : .994867 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern
```

Von OLS* verwendete Daten fuer :

```

y1 : -3.0000 -3.0000 11.0000 0.0000 0.0000 -1.0000 6.0000 -5.0000
y2 : -5.0000 -17.0000 25.0000 -5.0000 5.0000 0.0000 17.0000 -20.0000

x1 : -2.0000 -1.0000 0.0000 0.0000 0.0000 0.0000 1.0000 2.0000
x2 : -2.0000 -2.0000 2.0000 -2.0000 2.0000 0.0000 2.0000 0.0000
x3 : 0.0000 -1.0000 2.0000 0.0000 0.0000 0.0000 1.0000 -2.0000
o3 : 0.0000 -2.0000 2.0000 0.0000 0.0000 2.0000 0.0000 -2.0000

```

----- (1)

y1 = 2.0000 x1 5.0000 x3
(5.1968) (15.4919)

s**2 : 1.000 () d : 2.333 () B(adj.) : .9727 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 2 , 4 Freiheitsgraden : F* : .00

----- (2)

y2 = 3.0000 x2 10.0000 x3
(15.4143) (33.1563)

s**2 : .667 () d : 1.750 () B(adj.) : .9973 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 2 , 4 Freiheitsgraden : F* : .00

Korrelationsmatrix der U :

```

1.0000 0.0000
0.0000 1.0000

```

Prognosen aus Strukturform fuer :

```

y1 : -4.0000 -7.0000 10.0000 0.0000 0.0000 0.0000 7.0000 -5.0000
y2 : -5.0000 -15.0000 25.0000 -5.0000 5.0000 0.0000 15.0000 -20.0000

```

Prognosefehler aus Strukturform (=U) fuer :

```

y1 : 1.0000 -1.0000 1.0000 0.0000 0.0000 -1.0000 -1.0000 1.0000
y2 : 1.0000 -1.0000 -1.0000 0.0000 0.0000 0.0000 1.0000 .0000

```

----- (3)

y1 = 1.6154 x1 3.0769 o3
(1.2651) (3.0073)

s**2 : 16.385 () d : 2.803 () B(adj.) : .5520 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 2 , 4 Freiheitsgraden : F* : 30.77

----- (4)

y2 = 4.5000 x2 5.5000 o3
(2.6973) (2.6917)

s**2 : 55.667 () d : 2.272 () B(adj.) : .7789 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 2 , 4 Freiheitsgraden : F* : 164.99

Korrelationsmatrix der U :

```

1.0000 .0560
.0560 1.0000

```

y1 = 1.0000 x1 (.5334) (5)

s**2 : 35.143 () d : 2.691 () B(adj.) : .0391 () Fr.-grade : 7 () "t"-Werte : Werte in Klammern

Joint-Test auf Unabh angigkeit mit 3 , 4 Freiheitsgraden : F* : 53.33

y2 = 6.3333 x2 (3.0231) (6)

s**2 : 105.333 () d : 1.745 () B(adj.) : .5653 () Fr.-grade : 7 () "t"-Werte : Werte in Klammern

Joint-Test auf Unabh angigkeit mit 3 , 4 Freiheitsgraden : F* : 244.47

Korrelationsmatrix der U :

1.0000 .7514
.7514 1.0000

Fuer p3 sind Instrumentalvariable: x2

y1 = 2.0000 x1 5.0000 p3 (6.0960) (11.4415) (7)

s**2 : 1.0000 () B(adj.) : .972656 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern

Fuer p3 sind Instrumentalvariable: y2

y1 = 2.0000 x1 5.0000 p3 (6.1910) (15.1414) (8)

s**2 : 1.0000 () B(adj.) : .972656 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern

Fuer p3 sind Instrumentalvariable: y2 x2

y1 = 2.0000 x1 5.0000 p3 (6.1963) (15.4598) (9)

s**2 : 1.0000 () B(adj.) : .972656 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern

Fuer p3 sind Instrumentalvariable: x1

y2 = 3.0000 x2 10.0000 p3 (13.0658) (21.1061) (10)

s**2 : .6666 () B(adj.) : .997255 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern

Fuer p3 sind Instrumentalvariable: y1

y2 = 3.0000 x2 10.0000 p3 (14.9625) (29.9037) (11)

s**2 : .6666 () B(adj.) : .997255 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern

Fuer p3 sind Instrumentalvariable: y1 x1

y2 = 3.0000 x2 10.0000 p3 (15.3485) (32.6410) (12)

s**2 : .6666 () B(adj.) : .997255 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern

Gamma OLS fuer :

$$y1 = 1.6154 x1 + 3.0769 p3$$

$$(1.1227) \quad (.8875)$$

(3')

s**2 : 12.288 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

$$y2 = 4.5000 x2 + 5.5000 p3$$

$$(1.4448) \quad (1.7595)$$

(4')

s**2 : 41.750 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

R = Korrelationsmatrix der Residuen :

$$\begin{pmatrix} 1.0000 & .8660 \\ .8660 & 1.0000 \end{pmatrix}$$

Inverse von R :

$$\begin{pmatrix} 3.9994 & -3.4635 \\ -3.4635 & 3.9994 \end{pmatrix}$$

Gamma des SUR-Schaetzers nach 0 Iterationen fuer:

$$y1 = 2.1570 x1 + 3.1446 p3$$

$$(.6766) \quad (.8804)$$

(13)

s**2 : 12.646 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

$$y2 = 3.4156 x2 + 6.0422 p3$$

$$(.8707) \quad (1.6730)$$

(14)

s**2 : 44.690 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

$$\begin{matrix} y1 : & 1.3141 & .4463 & 4.7107 & 0.0000 & 0.0000 & -7.2893 & 3.8430 & -3.0248 \\ y2 : & 1.8312 & 1.9156 & 6.0844 & .8312 & -.8312 & -12.0844 & 10.1668 & -7.9156 \end{matrix}$$

R der SURE-Residuen :

$$\begin{pmatrix} 1.0000 & .9624 \\ .9624 & 1.0000 \end{pmatrix}$$

Maximale rel. Abweichung : .1653737E+00 nach 0 Iterationen

Gamma des SUR-Schaetzers nach 8 Iterationen fuer:

$$y1 = 2.0929 x1 + 3.1366 p3$$

$$(.3852) \quad (.8875)$$

(15)

s**2 : 12.566 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

$$y2 = 3.1936 x2 + 6.1532 p3$$

$$(.5146) \quad (1.7153)$$

(16)

s**2 : 46.017 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

$$\begin{matrix} y1 : & 1.1858 & .3661 & 4.7268 & 0.0000 & 0.0000 & -7.2732 & 3.9071 & -2.9126 \\ y2 : & 1.3871 & 1.6936 & 6.3064 & .3871 & -.3871 & -12.3064 & 10.6129 & -7.6936 \end{matrix}$$

R der SURE-Residuen :

$$\begin{pmatrix} 1.0000 & .9640 \\ .9640 & 1.0000 \end{pmatrix}$$

Maximale rel. Abweichung : .2735631E-05 nach 8 Iterationen

Von OLS* verwendete Daten fuer :

```

y1 : -7.0000 -5.0000 -1.0000 1.0000 1.0000 11.0000
y2 : 13.0000 7.0000 1.0000 -6.0000 -1.0000 -14.0000

x1 : -2.0000 -1.0000 0.0000 0.0000 1.0000 2.0000
x2 : 2.0000 2.0000 -1.0000 -1.0000 -1.0000 -1.0000
x3 : -4.0000 -2.0000 -2.0000 2.0000 0.0000 6.0000

```

$$y1 = 2.0000 x1 + 1.0000 x3 \quad (1)$$

(2.5495) (3.2249)

s**2 : 1.500 () d : 2.667 () B(adj.) : .9521 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern
 Joint-Test auf Unabhaengigkeit mit 1, 3 Freiheitsgraden : F* : .00

$$y2 = 2.0000 x2 - 2.0000 x3 \quad (2)$$

(5.2678) (-12.1655)

s**2 : 1.000 () d : 1.750 () B(adj.) : .9989 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern
 Joint-Test auf Unabhaengigkeit mit 1, 3 Freiheitsgraden : F* : .00

Korrelationsmatrix der U :

```

1.0000 .0000
.0000 1.0000

```

Prognosen aus Strukturform fuer :

```

y1 : -8.0000 -4.0000 -2.0000 2.0000 2.0000 10.0000
y2 : 12.0000 8.0000 2.0000 -6.0000 -2.0000 -14.0000

```

Prognosefehler aus Strukturform (=U) fuer :

```

y1 : 1.0000 -1.0000 1.0000 -1.0000 -1.0000 1.0000
y2 : 1.0000 -1.0000 -1.0000 .0000 1.0000 .0000

```

Gamma OLS fuer :

$$y1 = 4.2000 x1 \quad (3)$$

(.6000)

s**2 : 3.600 () Fr.-grade : 5 !! () "Sigma"-Werte : Werte in Klammern

$$y2 = 5.0000 x2 \quad (4)$$

(1.4530)

s**2 : 25.333 () Fr.-grade : 5 !! () "Sigma"-Werte : Werte in Klammern

R = Korrelationsmatrix der Residuen :

```

1.0000 -.6388
-.6388 1.0000

```

Inverse von R :

```

1.6892 1.0790
1.0790 1.6892

```

Gamma des SUR-Schaetzers nach 0 Iterationen fuer:

$$y1 = 3.7111 \times 1 \quad (5)$$

(.5423)

s**2 : 3.998 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

$$y2 = 4.6328 \times 2 \quad (6)$$

(1.3133)

s**2 : 25.683 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

y1 :	.4222	-1.2689	-1.0000	1.0000	-2.7111	3.5778
y2 :	3.7343	-2.2657	5.6328	-1.3672	3.6328	-9.3672

R der SUR-Residuen :

1.0000	-.7555
-.7555	1.0000

Maximale rel. Abweichung : .2531722E+00 nach 0 Iterationen

Gamma des SUR-Schaetzers nach 34 Iterationen fuer:

$$y1 = 3.8774 \times 1 \quad (7)$$

(.5353)

s**2 : 5.701 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

$$y2 = 3.6459 \times 2 \quad (8)$$

(1.1023)

s**2 : 29.000 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

y1 :	-.8453	-1.9226	-1.0000	1.0000	-2.8774	4.8453
y2 :	5.7081	-.2919	4.6459	-2.3541	2.6459	-10.3541

R der SUR-Residuen :

1.0000	-.8676
-.8676	1.0000

Maximale rel. Abweichung : .6105142E-05 nach 34 Iterationen
===== (1)

y1 = 2.1659 x1 1.0000"ZZ" (9)
 (2.6604) (2.9079)

s**2 : 1.7341 () d : 2.489 () B(adj.) : .955208 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern

y2 = 1.7333 x2 1.0000"ZZ" (10)
 (1.8603) (4.9861)

s**2 : 5.2667 () d : 3.015 () B(adj.) : .941740 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern

Korrelationsmatrix der Residuen :

1.0000 .8161
 .8161 1.0000

Korrelationsmatrix der "Hilfsvariablen" :

1.0000 -.9566
 -.9566 1.0000

Prognosefehler fuer :

y1 : 1.4878 -1.4878 .4341 -.8439 -.6683 1.0781
 y2 : 2.2444 -2.2444 1.3111 -2.8444 .8222 .7111

Werte der "Hilfsvariablen" fuer :

y1 : -4.1561 -1.3463 -1.4341 1.8439 -.4976 5.5902
 y2 : 7.2889 5.7778 1.4222 -1.4222 -.0889 -12.9778

y1 = -.4683 y2 2.1659 x1 .9659 x2 (11)
 (-2.4459) (1.6541) (1.2117)

s**2 : 2.312 () d : 2.489 () B(adj.) : .9416 () Fr.-grade : 3 () "t"-Werte : Werte in Klammern

y2 = -1.4222 y1 1.3333 x1 1.7333 x2 (12)
 (-2.4459) (.4357) (1.2667)

s**2 : 7.022 () d : 3.015 () B(adj.) : .9223 () Fr.-grade : 3 () "t"-Werte : Werte in Klammern

Von OLS* verwendete Daten fuer :

```
y1 : -3.0000 -8.0000 11.0000 -1.0000 6.0000 -5.0000
y2 : 1.0000 -14.0000 22.0000 3.0000 11.0000 -23.0000

x1 : -2.0000 -1.0000 0.0000 0.0000 1.0000 2.0000
x2 : 0.0000 -1.0000 1.0000 1.0000 0.0000 -1.0000
x3 : 0.0000 -1.0000 2.0000 0.0000 1.0000 -2.0000
```

y1 = 2.0000 x1 5.0000 x3 (VIII.1)
(5.0596) (12.6491)

s**2 : 1.500 () d : 2.667 () B(adj.) : .9707 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 1, 3 Freiheitsgraden : F* : 0.00

y2 = 3.0000 x2 10.0000 x3 (VIII.2)
(3.6742) (19.3649)

s**2 : 1.000 () d : 1.750 () B(adj.) : .9963 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 1, 3 Freiheitsgraden : F* : .00

Korrelationsmatrix der U :

```
1.0000 0.0000
0.0000 1.0000
```

Prognosen aus Strukturform fuer :

```
y1 : -4.0000 -7.0000 10.0000 0.0000 7.0000 -6.0000
y2 : 0.0000 -13.0000 23.0000 3.0000 10.0000 -23.0000
```

Prognosefehler aus Strukturform (=U) fuer :

```
y1 : 1.0000 -1.0000 1.0000 -1.0000 -1.0000 1.0000
y2 : 1.0000 -1.0000 -1.0000 0.0000 1.0000 0.0000
```

y1 = 1.9959 x1 1.0000"ZZ" (VIII.3)
(4.6786) (11.6956)

s**2 : 1.7473 () d : 2.548 () B(adj.) : .965873 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern

y2 = 3.7512 x2 1.0000"ZZ" (VIII.4)
(1.8182) (7.2891)

s**2 : 6.6338 () d : 2.525 () B(adj.) : .975247 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern

Korrelationsmatrix der Residuen :

```
1.0000 -.9638
-.9638 1.0000
```

Korrelationsmatrix der "Hilfsvariablen" :

```
1.0000 .9963
.9963 1.0000
```

Werte der "Hilfsvariablen" fuer :

```
y1 : .4946 -5.5070 3.4641 -.0660 5.4411 -9.9588
y2 : 1.9718 -11.2207 20.6573 -1.8779 7.4648 -16.9583
```

Gamma OLS fuer :

$$y1 = 1.0000 x1 \quad (VII.5)$$

s**2 : 41.000 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

$$y2 = 15.5000 x2 \quad (VIII.6)$$

s**2 : 63.167 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

R = Korrelationsmatrix der Residuen :

1.0000 .5895
.5895 1.0000

Inverse von R :

1.5326 -.9035
-.9035 1.5326

Gamma des SUR-Schaetzers nach 0 Iterationen fuer:

$$y1 = 1.5697 x1 \quad (VIII.7)$$

s**2 : 41.541 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

$$y2 = 11.0056 x2 \quad (VIII.8)$$

s**2 : 76.633 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

y1 : .1393 -6.4303 11.0000 -1.0000 4.4303 -8.1393
y2 : 1.0000 -2.9944 10.9944 -8.0056 11.0000 -11.9944

R der SURE-Residuen :

1.0000 .8705
.8705 1.0000

Maximale rel. Abweichung : .6372194E+00 nach 0 Iterationen

Gamma des SUR-Schaetzers nach 23 Iterationen fuer:

$$y1 = 1.9946 x1 \quad (VIII.9)$$

s**2 : 42.649 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

$$y2 = 3.7523 x2 \quad (VIII.10)$$

s**2 : 155.172 () Fr.-grade : 6 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

y1 : .9892 -6.0054 11.0000 -1.0000 4.0054 -8.9892
y2 : 1.0000 -10.2477 18.2477 -.7523 11.0000 -13.2477

R der SURE-Residuen :

1.0000 .9856
.9856 1.0000

Maximale rel. Abweichung : .6096785E-05 nach 23 Iterationen

Von OLS* verwendete Daten fuer :

y1 :	-5.0000	-6.0000	4.0000	-3.0000	3.0000	-1.0000	4.0000	5.0000
y2 :	-4.0000	-9.0000	10.0000	-5.0000	5.0000	0.0000	9.0000	-6.0000
y3 :	-6.0000	-7.0000	6.0000	-2.0000	2.0000	2.0000	5.0000	0.0000
x1 :	-2.0000	-1.0000	0.0000	0.0000	0.0000	0.0000	1.0000	2.0000
x2 :	0.0000	-1.0000	2.0000	0.0000	0.0000	0.0000	1.0000	-2.0000
x3 :	-1.0000	-1.0000	1.0000	-1.0000	1.0000	0.0000	1.0000	0.0000

y1 = 2.0000 x1 3.0000 x3 (IX.1)
(5.4160) (6.2929)

s**2 : 1.000 () d : 2.333 () B(adj.) : .9527 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 1, 5 Freiheitsgraden : F* : .00

y2 = 3.0000 x2 5.0000 x3 (IX.2)
(9.9499) (12.8452)

s**2 : .667 () d : 1.750 () B(adj.) : .9872 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 1, 5 Freiheitsgraden : F* : .00

y3 = 2.0000 x1 2.0000 x2 2.0000 x3 (IX.3)
(3.8139) (3.8139) (2.5820)

s**2 : 1.200 () d : 2.667 () B(adj.) : .9468 () Fr.-grade : 5 () "t"-Werte : Werte in Klammern

Korrelationsmatrix der U :

1.0000	-.0000	0.0000
-.0000	1.0000	0.0000
0.0000	0.0000	1.0000

Prognosen aus Strukturform fuer :

y1 :	-7.0000	-5.0000	3.0000	-3.0000	3.0000	0.0000	5.0000	4.0000
y2 :	-5.0000	-8.0000	11.0000	-5.0000	5.0000	0.0000	8.0000	-6.0000
y3 :	-6.0000	-6.0000	6.0000	-2.0000	2.0000	0.0000	6.0000	0.0000

Prognosefehler aus Strukturform (=U) fuer :

y1 :	1.0000	-1.0000	1.0000	-.0000	.0000	-1.0000	-1.0000	1.0000
y2 :	1.0000	-1.0000	-1.0000	-.0000	0.0000	0.0000	1.0000	0.0000
y3 :	0.0000	-1.0000	0.0000	0.0000	0.0000	2.0000	-1.0000	0.0000

Gamma OLS fuer :

$$y1 = 3.2000 \cdot x1 \quad (IX.4)$$

(.7550)

s**2 : 5.700 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

$$y2 = 5.0000 \cdot x2 \quad (IX.5)$$

(1.1937)

s**2 : 14.250 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

$$y3 = 3.0000 \cdot x1 + 3.0000 \cdot x2 \quad (IX.6)$$

(.4270) (.4270)

s**2 : 1.750 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

R = Korrelationsmatrix der Residuen :

1.0000	.5160	.4749
.5160	1.0000	.5006
.4749	.5006	1.0000

Inverse von R :

1.4899	-.5531	-.4307
-.5531	1.5398	-.5082
-.4307	-.5082	1.4590

Gamma des SUR-Schaetzers nach 0 Iterationen fuer:

$$y1 = 2.4859 \cdot x1 \quad (IX.7)$$

(.6502)

s**2 : 6.337 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

$$y2 = 3.9418 \cdot x2 \quad (IX.8)$$

(1.0281)

s**2 : 15.650 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

$$y3 = 2.5290 \cdot x1 + 2.5644 \cdot x2 \quad (IX.9)$$

(.3631) (.3751)

s**2 : 2.162 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

y1 :	-1.0282	-3.5141	4.0000	-3.0000	3.0000	-1.0000	1.5141	.0282
y2 :	-4.0000	-5.0582	2.1165	-5.0000	5.0000	0.0000	5.0582	1.8635
y3 :	-.9421	-1.9066	.8711	-2.0000	2.0000	2.0000	-.0934	.0709

R der SURE-Residuen :

1.0000	.8543	.7096
.8543	1.0000	.7504
.7096	.7504	1.0000

Maximale rel. Abweichung : .8297148E+00 nach 0 Iterationen

Gamma des SUR-Schaetzers nach 23 Iterationen fuer:

y1 = 1.9294 x1 (.3171) (IX.10)

s**2 : 7.718 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

y2 = 3.3164 x2 (.4814) (IX.11)

s**2 : 17.793 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

y3 = 2.0521 x1 2.1565 x2 (.2931) (.3251) (IX.12)

s**2 : 3.363 () Fr.-grade : 8 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

y1 :	-2.1413	-4.0706	4.0000	-3.0000	3.0000	-1.0000	2.0706	1.1413
y2 :	-4.0000	-5.6836	3.3672	-5.0000	5.0000	0.0000	5.6836	.6328
y3 :	-1.8959	-2.7915	1.6870	-2.0000	2.0000	2.0000	.7915	.2089

R der SURE-Residuen :

1.0000	.9350	.8354
.9350	1.0000	.8687
.8354	.8687	1.0000

Maximale rel. Abweichung : .6168177E-05 nach 23 Iterationen

y1 = 2.1111 x1 1.0000 x2 (5.2763) (5.5955) (IX.13)

s**2 : 1.2222 () d : 2.247 () B(adj.) : .942192 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern

y2 = 3.6250 x2 1.0000 x1 (6.3264) (5.9544) (IX.14)

s**2 : 2.7500 () d : 1.636 () B(adj.) : .947115 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern

y2 = 1.8182 x1 1.1515 x2 1.0000 x3 (2.9646) (1.3526) (2.4072) (IX.15)

s**2 : 1.2970 () d : 2.588 () B(adj.) : .942539 () Fr.-grade : 5 () "t"-Werte : Werte in Klammern

Korrelationsmatrix der Residuen :

1.0000	-.8333	.0000
-.8333	1.0000	-.0000
.0000	-.0000	1.0000

Korrelationsmatrix der "Hilfsvariablen" :

1.0000	.9291	.9604
.9291	1.0000	.8868
.9604	.8868	1.0000

Werte der "Hilfsvariablen" fuer :

y1 :	-2.2222	-3.4444	2.4444	-2.7778	2.7778	0.0000	3.4444	-2.2222
y2 :	-3.7500	-5.6250	5.0000	-3.7500	3.7500	-1.2500	3.1250	2.5000
y3 :	-1.9394	-3.4545	3.5151	-1.8788	1.8788	-.1212	3.2121	-1.2121

Von OLS/EIV verwendete Daten fuer :

y1 :	-6.0000	-6.0000	4.0000	-3.0000	3.0000	-1.0000	4.0000	5.0000
y2 :	-4.0000	-9.0000	10.0000	-5.0000	5.0000	0.0000	9.0000	-6.0000
y3 :	-6.0000	-7.0000	6.0000	-2.0000	2.0000	2.0000	5.0000	0.0000
x1 :	-2.0000	-1.0000	0.0000	0.0000	0.0000	0.0000	1.0000	2.0000
x2 :	0.0000	-1.0000	2.0000	0.0000	0.0000	0.0000	1.0000	-2.0000

Fuer y2 sind Instrumentalvariable: y3 x1 x2

y1 = .5000 y2 2.0000 x1 -1.0000 x2 (2.9277) (3.2129) (-1.5237) (IX.16)

s**2 : 1.2000 () B(adj.) : .943244 () Fr.-grade : 5 () "t"-Werte : Werte in Klammern

Fuer y1 sind Instrumentalvariable: y3 x1 x2

y2 = 1.6667 y1 -3.3333 x1 3.0000 x2 (5.9763) (-3.2748) (5.9025) (IX.17)

s**2 : .8000 () B(adj.) : .984616 () Fr.-grade : 5 () "t"-Werte : Werte in Klammern

Fuer y1 sind Instrumentalvariable: y2 x1 x2

y3 = .6667 y1 .6667 x1 2.0000 x2 (2.4845) (.6643) (3.7333) (IX.18)

s**2 : 1.2000 () B(adj.) : .946836 () Fr.-grade : 5 () "t"-Werte : Werte in Klammern

Fuer y2 sind Instrumentalvariable: y1 x1 x2

y3 = .4000 y2 2.0000 x1 .8000 x2 (2.2361) (3.5083) (.7652) (IX.19)

s**2 : 1.2000 () B(adj.) : .946837 () Fr.-grade : 5 () "t"-Werte : Werte in Klammern

Fuer y2 sind Instrumentalvariable: y3

y1 = .6000 y2 2.0000 x1 -1.0000 x2 (2.9277) (3.2128) (-1.5237) (IX.20)

s**2 : 1.2000 () B(adj.) : .943242 () Fr.-grade : 5 () "t"-Werte : Werte in Klammern

Fuer y2 sind Instrumentalvariable: y3

y1 = .3094 y2 2.7668 x1 (3.9264) (6.0516) (IX.21)

s**2 : 1.9686 () B(adj.) : .906890 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern

Fuer y2 sind Instrumentalvariable: y3 x1 x2

y1 = .3020 y2 2.7771 x1 (3.7200) (5.8782) (IX.22)

s**2 : 2.1028 () B(adj.) : .900542 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern

$$- A \cdot X \cdot 1 -$$

Von OLS* verwendete Daten fuer :

I :	33.1000	45.0000	77.2000	44.6000	48.1000	74.4000	113.0000	91.9000	61.3000	56.8000
I' :	93.6000	159.9000	147.2000	146.3000	98.3000	93.5000	135.2000	157.3000	179.5000	189.6000
	12.9300	25.9000	35.0500	22.8900	18.8400	28.5700	48.5100	43.3400	37.0200	37.810
	39.2700	53.4600	55.5600	49.5600	32.0400	32.2400	54.3800	71.7600	90.0800	68.600
Kon :	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.000
	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.000
F-1 :	1170.6000	2015.8000	2803.3000	2039.7000	2256.2000	2132.2000	1834.1000	1588.0000	1749.4000	1687.200
	2007.7000	2208.3000	1656.7000	1604.4000	1431.8000	1610.5000	1819.4000	2079.7000	2371.6000	2759.900
F'-1 :	191.5000	516.0000	729.0000	560.4000	519.9000	628.5000	537.1000	561.2000	617.2000	626.700
	737.2000	760.5000	581.4000	662.3000	583.8000	635.2000	723.6000	864.1000	1193.5000	1198.900
C-1 :	97.8000	104.4000	118.0000	156.2000	172.6000	186.6000	220.9000	287.0000	319.9000	321.300
	319.6000	346.0000	456.4000	543.4000	618.3000	647.4000	671.3000	726.1000	800.3000	888.900
C'-1 :	1.0000	.0000	7.4000	18.1000	23.5000	26.5000	36.2000	60.8000	84.4000	91.200
	92.4000	86.0000	111.1000	130.6000	141.8000	136.7000	129.7000	145.5000	174.8000	213.500

$$I = -9.9561 \text{ Kon} \quad .0266 \text{ F-1} \quad .1517 \text{ C-1} \quad (X.1)$$

$$(-.3173) \quad (1.7057) \quad (5.9015)$$

s**2 : 777.446 () d : 1.072 () B(adj.) : .6706 () Fr.-grade : 17 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 2,15 Freiheitsgraden : F* : .53

$$I' = -.5094 \text{ Kon} \quad .0529 \text{ F'-1} \quad .0924 \text{ C'-1} \quad (X.2)$$

$$(-.0636) \quad (3.3677) \quad (1.6472)$$

s**2 : 104.328 () d : 1.413 () B(adj.) : .7144 () Fr.-grade : 17 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 2,15 Freiheitsgraden : F* : 2.49

Korrelationsmatrix der U :

1.0000	.7290
.7290	1.0000

Prognosen aus Strukturform fuer :

I :	35.9603	59.4025	82.3746	67.8949	76.1310	74.9523	72.2505	75.8646	85.0193	83.580
	91.8320	101.1629	103.2643	115.0730	121.8521	131.0111	140.1832	155.4072	174.4132	198.163
I' :	9.7852	26.8579	38.7342	30.8050	29.1618	35.1833	31.2451	34.7931	39.9360	41.067
	47.0225	47.6635	40.5095	46.5907	43.4734	45.7209	49.7605	58.6416	78.7724	82.105

Prognosefehler aus Strukturform (=U) fuer :

I :	-2.8603	-14.4025	-5.1746	-23.2949	-28.0310	-5.5623	40.7495	16.0354	-23.7193	-26.780
	1.7690	58.7371	43.9357	31.2270	-23.5521	-37.5111	-4.9832	1.8928	5.0868	-8.563
I' :	3.1438	-9.9579	-3.6842	-7.9150	-10.3218	-6.6133	17.2649	8.5469	-2.9160	-3.257
	-7.7525	5.7965	15.0504	2.9693	-11.4334	-13.4809	4.6195	13.1384	11.3076	-13.505

OLS-Residuum von (X.2) ueber jenem von (X.1)

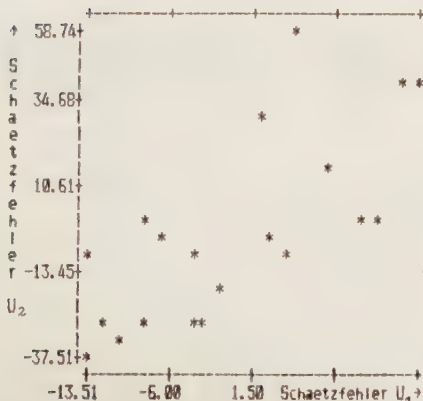


Abb. (X.A)

Gamma des SUR-Schaetzers nach 0 Iterationen fuer:

$$I = -27.7189 \text{ Kon } .0383 \text{ F-1 } .1390 \text{ C-1} \quad (X.3)$$

(27.0327) (.0133) (.0230)

s**2 : 689.417 () Fr.-grade : 20 !! () "Sigma"-Werte : Werte in Klammern

$$I' = -1.2520 \text{ Kon } .0576 \text{ F'-1 } .0640 \text{ C'-1} \quad (X.4)$$

(6.9563) (.0134) (.0489)

s**2 : 90.065 () Fr.-grade : 20 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

I :	2.3756	-19.0216	-18.8815	-27.5393	-34.6136	-5.5097	39.7415	18.7680	-22.4783	-24.790
I' :	-.0320	54.9125	47.9945	37.0019	-14.7996	-30.4916	-1.1175	4.3912	5.0920	-12.002
	3.0507	-2.6361	-6.1834	-9.3117	-11.3732	-8.0937	16.4931	8.3603	-2.6969	-2.889
	-7.8743	5.3824	16.1980	4.2882	-9.4245	-11.8603	5.6215	13.9252	11.3674	-12.323

R der SURE-Residuen :

1.0000 .7650
.7650 1.0000

Maximale rel. Abweichung : .8039967E-01 nach 0 Iterationen

Gamma des SUR-Schaetzers nach 7 Iterationen fuer:

$$I = -30.7461 \text{ Kon } .0405 \text{ F-1 } .1359 \text{ C-1} \quad (X.5)$$

(27.3456) (.0134) (.0235)

s**2 : 702.228 () Fr.-grade : 20 !! () "Sigma"-Werte : Werte in Klammern

$$I' = -1.7012 \text{ Kon } .0594 \text{ F'-1 } .0557 \text{ C'-1} \quad (X.6)$$

(6.9284) (.0133) (.0486)

s**2 : 90.953 () Fr.-grade : 20 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

I :	3.1317	-20.1040	-21.6538	-28.5134	-36.0130	-6.5928	39.4206	19.1961	-22.3055	-24.476
I' :	-.4283	54.1569	48.7951	38.1878	-13.0016	-28.9962	.9926	5.0990	5.3892	-12.285
	3.1632	-3.0695	-6.9281	-9.6780	-11.6252	-8.5000	16.3160	8.3445	-2.6146	-2.767
	-7.9326	5.2312	16.5619	4.6735	-8.8117	-11.3781	5.8936	14.0860	11.2026	-12.161

R der SURE-Residuen :

1.0000 .7730
.7730 1.0000

Maximale rel. Abweichung : .1405979E-05 nach 7 Iterationen

Abb. (X.B)

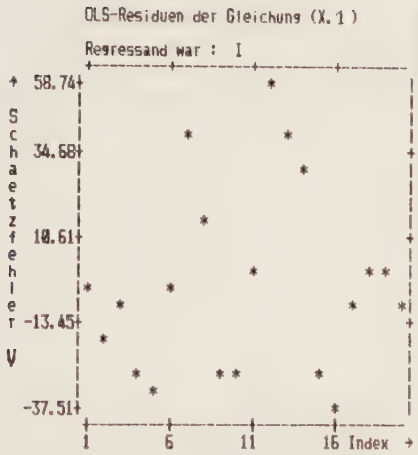


Abb. (X.C)

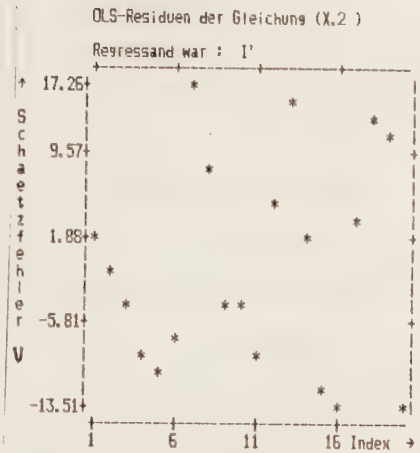


Abb. (X.D)

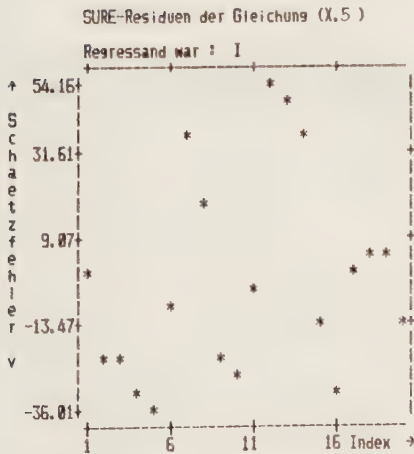
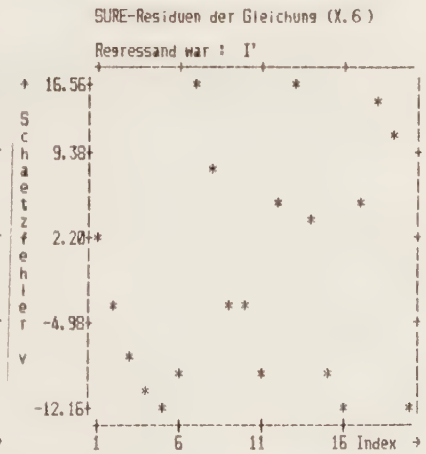


Abb. (X.E)



Von SURE verwendete Daten fuer : (t = 1 - 10)

I a :	33.1000	45.0000	77.2000	44.6000	48.1000	74.4000	113.0000	91.9000	61.3000	56.800
I'a :	12.9300	25.9000	35.0500	22.8900	18.8400	28.5700	48.5100	43.3400	37.0200	37.810
Kon :	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.000
F a :	1170.6000	2015.8000	2803.3000	2039.7000	2256.2000	2132.2000	1834.1000	1588.0000	1749.4000	1687.200
F'a :	191.5000	516.0000	729.0000	560.4000	519.9000	628.5000	537.1000	561.2000	617.2000	626.700
C a :	97.8000	104.4000	118.0000	156.2000	172.6000	186.6000	220.9000	287.8000	319.9000	321.300
C'a :	1.8000	.8000	7.4000	18.1000	23.5000	26.5000	36.2000	60.6000	84.4000	91.200

Gamma DLS fuer :

I a = 3.6248 Kon .0175 F a .1371 C a
(40.2184) (.0163) (.0832)

(X.9)

s**2 : 417.645 () Fr.-grade : 10 !! () "Sigma"-Werte : Werte in Klammern

I'a = 7.5492 Kon .0134 F'a .1493 C'a
(10.0021) (.0190) (.0812)

(X.10)

s**2 : 55.730 () Fr.-grade : 10 !! () "Sigma"-Werte : Werte in Klammern

R = Korrelationsmatrix der Residuen :

1.0000 .9218
.9218 1.0000

Inverse von R :

6.6559 -6.1355
-6.1355 6.6559

Gamma des SUR-Schaetzers nach 0 Iterationen fuer:

I a = -15.7169 Kon .0242 F a .1692 C a
(33.8586) (.0131) (.0772)

(X.11)

s**2 : 427.885 () Fr.-grade : 10 !! () "Sigma"-Werte : Werte in Klammern

I'a = 5.2497 Kon .0348 F'a .1914 C'a
(8.3579) (.0153) (.0752)

(X.12)

s**2 : 57.703 () Fr.-grade : 10 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

I a :	3.9327	-5.7428	5.0939	-15.4865	-20.0022	6.9304	46.9421	20.4788	-19.4597	-22.691
I'a :	.6622	2.5152	2.9791	-5.3536	-9.0259	-3.6547	17.6136	6.8948	-5.8942	-6.737

R der SURE-Residuen :

1.0000 .9400
.9400 1.0000

Maximale rel. Abweichung : .5026258E-01 nach 0 Iterationen

Gamma des SUR-Schaetzers nach 5 Iterationen fuer:

$$I\ a = -19.7146\ Kon\ .0249\ F\ a\ .1826\ C\ a \quad (X.13)$$

$$(32.7746) \quad (.0125) \quad (.0767)$$

s**2 : 433.980 () Fr.-grade : 10 !! () "Sigma"-Werte : Werte in Klammern

$$I'\ a = 4.3962\ Kon\ .0357\ F'\ a\ .2024\ C'\ a \quad (X.14)$$

$$(8.1532) \quad (.0147) \quad (.0728)$$

s**2 : 58.974 () Fr.-grade : 10 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

I a :	5.8072	-4.5467	5.5585	-14.9986	-19.8844	6.9479	46.7899	19.5254	-20.9543	-24.161
I' a :	1.3322	2.9185	3.1279	-5.1777	-8.8744	-3.6290	17.6114	6.6027	-6.4925	-7.418

R der SURE-Residuen :

1.0000 .9423
.9423 1.0000

Maximale rel. Abweichung : .9919080E-05 nach 6 Iterationen

Von SURE verwendete Daten fuer : (t = 11 - 20)

I b :	93.6000	159.9000	147.2000	146.3000	98.3000	93.5000	135.2000	157.3000	179.5000	189.600
I' b :	39.2700	53.4600	55.5600	49.5600	32.0400	32.2400	54.3600	71.7000	90.0000	68.600
KON :	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.000
F b :	2007.7000	2208.3000	1656.7000	1604.4000	1431.8000	1510.5000	1819.4000	2079.7000	2371.6000	2759.900
F' b :	737.2000	760.5000	581.4000	662.3000	583.8000	635.2000	723.9000	854.1000	1193.5000	1188.900
C b :	319.6000	346.0000	456.4000	543.4000	618.3000	647.4000	671.3000	726.1000	800.3000	898.900
C' b :	92.4000	85.0000	111.1000	130.6000	141.8000	136.7000	129.7000	145.5000	174.9000	213.500

Gamma OLS fuer :

$$I\ b = 8.5703\ KON\ .0553\ F\ b\ .0389\ C\ b \quad (X.15)$$

$$(35.8487) \quad (.0189) \quad (.0416)$$

s**2 : 454.214 () Fr.-grade : 10 !! () "Sigma"-Werte : Werte in Klammern

$$I'\ b = 5.5594\ KON\ .0766\ F'\ b\ -.0852\ C'\ b \quad (X.16)$$

$$(12.8419) \quad (.0211) \quad (.1268)$$

s**2 : 95.233 () Fr.-grade : 10 !! () "Sigma"-Werte : Werte in Klammern

R = Korrelationsmatrix der Residuen :

1.0000 .8015
.8015 1.0000

Inverse von R :

2.7969 -2.2419
-2.2419 2.7969

Gamma des SUR-Schaetzers nach 0 Iterationen fuer:

$$I' b = \begin{pmatrix} 3.6057 \\ 33.9424 \end{pmatrix} \begin{pmatrix} .0597 \\ .0163 \end{pmatrix} F' b \begin{pmatrix} .0002 \\ .0364 \end{pmatrix} C' b \quad (X.17)$$

s**2 : 501.784 () Fr.-grade : 10 !! () "Sigma"-Werte : Werte in Klammern

$$I' b = \begin{pmatrix} 8.2416 \\ 12.3219 \end{pmatrix} \begin{pmatrix} .0861 \\ .0160 \end{pmatrix} F' b \begin{pmatrix} -.1590 \\ .1072 \end{pmatrix} C' b \quad (X.18)$$

s**2 : 98.476 () Fr.-grade : 10 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

I' b :	-50.0694	2.2593	27.9897	30.7158	-5.2683	-22.5334	4.5973	8.5375	10.3701	-6.621
I' b :	-17.5797	-6.4141	15.1030	5.2349	-3.7429	-8.7810	4.6146	12.4419	7.0275	-7.904

R der SURE-Residuen :

1.0000	.8533
.8533	1.0000

Maximale rel. Abweichung : .1512054E+00 nach 0 Iterationen

Gamma des SUR-Schaetzers nach 13 Iterationen fuer:

$$I' b = \begin{pmatrix} 3.5439 \\ 35.9236 \end{pmatrix} \begin{pmatrix} .0761 \\ .0162 \end{pmatrix} F' b \begin{pmatrix} -.0204 \\ .0366 \end{pmatrix} C' b \quad (X.19)$$

s**2 : 561.745 () Fr.-grade : 10 !! () "Sigma"-Werte : Werte in Klammern

$$I' b = \begin{pmatrix} 10.2525 \\ 12.3892 \end{pmatrix} \begin{pmatrix} .0922 \\ .0165 \end{pmatrix} F' b \begin{pmatrix} -.2107 \\ .0984 \end{pmatrix} C' b \quad (X.20)$$

s**2 : 104.573 () Fr.-grade : 10 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

I' b :	-56.2030	-4.6293	26.8953	31.7488	-1.5903	-19.3951	6.8961	10.3059	11.8067	-5.834
I' b :	-19.5049	-8.8123	15.0944	5.7417	-2.1785	-7.7937	4.7000	12.4894	6.5827	-6.319

R der SURE-Residuen :

1.0000	.8894
.8894	1.0000

Maximale rel. Abweichung : .6512263E-05 nach 13 Iterationen

$$I\ a = -63.5922\ Kon\ .0248\ F\ a\ .5489\ C\ a\ 1.0000''ZZ'' \quad (X.21)$$

(-3.9448) (4.2805) (9.8361) (8.6614)

s**2 : 51.5485 () d : 2.522 () B(adj.) : .915077 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern

$$I\ b = -24.3128\ Kon\ .0779\ F\ b\ -.0940\ C\ b\ 1.0000''ZZ'' \quad (X.22)$$

(-1.0603) (6.2457) (-2.4464) (4.5918)

s**2 : 167.7836 () d : 2.959 () B(adj.) : .861330 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern

$$I'\ a = 23.6100\ Kon\ .0465\ F'\ a\ .4419\ C'\ a\ 1.0000''ZZ'' \quad (X.23)$$

(5.1382) (5.8969) (8.5981) (7.4117)

s**2 : 9.1460 () d : 2.455 () B(adj.) : .927150 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern

$$I'\ b = 12.5908\ Kon\ .1015\ F'\ b\ -.4207\ C'\ b\ 1.0000''ZZ'' \quad (X.24)$$

(1.8264) (8.4572) (-4.6366) (5.4725)

s**2 : 26.4912 () d : 3.059 () B(adj.) : .921092 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern

Korrelationsmatrix der Residuen :

1.0000	-.5586	-.9163	-.3127
-.5586	1.0000	-.4123	-.8809
-.9163	-.4123	1.0000	-.1386
-.3127	-.8809	-.1386	1.0000

Korrelationsmatrix der "Hilfsvariablen" :

1.0000	-.7037	.8284	-.7378
-.7037	1.0000	-.9324	.9714
.8284	-.9324	1.0000	-.8876
-.7378	.9714	-.8876	1.0000

Prognosen fuer :

I a :	28.3543	54.8330	74.2402	49.1510	50.5718	62.8362	111.7910	94.4534	64.3073	54.862
I b :	109.7090	143.4969	156.4309	130.5079	98.3790	94.1962	131.6336	168.6303	178.8176	188.597
I' a :	14.5255	22.0352	35.0671	22.3438	18.6522	34.3138	47.0585	42.4350	36.0910	38.338
I' b :	35.8731	59.3336	48.2755	53.7869	33.1893	34.8584	56.2115	66.2999	89.9178	69.225

Prognosefehler fuer :

I a :	4.7457	-9.8330	2.9598	-4.5510	-2.4718	11.5638	1.2000	-2.5524	-3.0873	1.938
I b :	-16.1090	16.4031	-9.2309	15.7921	-.0790	-.6962	3.5664	-11.3303	.6824	1.003
I' a :	-1.5955	3.8648	-.0171	.5462	.1078	-5.7436	1.4515	.9650	.9250	-.528
I' b :	3.3969	-5.8736	7.2845	-4.2269	-1.1493	-2.6184	-1.8315	5.4900	.1622	-.625

Werte der "Hilfsvariablen" fuer :

I a :	9.2220	11.1059	3.5083	-23.5995	-36.5519	-28.8946	8.6311	-39.3179	-91.0668	-99.757
I b :	7.6675	28.3103	94.5821	80.9081	69.2618	53.8943	77.3051	99.1751	93.5978	81.457
I' a :	-18.7844	-25.9220	-25.7109	-35.3224	-39.5169	-30.2309	-17.5225	-34.1368	-53.5131	-54.712
I' b :	-12.6585	5.7454	23.4205	28.9252	21.0048	15.3126	24.7303	27.2278	29.7441	25.797

Von SURE verwendete Daten fuer :

Ig :	33.1000	45.0000	77.2000	44.6000	48.1000	74.4000	113.0000	91.9000	61.3000	56.800
Iw :	93.6000	159.9000	147.2000	146.3000	98.3000	93.5000	135.2000	157.3000	179.5000	189.600
Iw :	12.9300	25.9000	35.0500	22.8900	18.8400	28.5700	48.5100	43.3400	37.0200	37.810
Is :	39.2700	53.4600	55.5600	49.5600	32.0400	32.2400	54.3800	71.7800	90.0000	68.600
Is :	20.9900	35.5300	46.3900	26.2300	23.0400	36.1600	47.2800	44.5600	36.1600	28.820
	25.8700	42.0300	42.0500	43.4500	40.5100	41.8800	58.8200	64.5500	64.1000	45.930
Kon :	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.000
	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.000
Fg :	1170.6000	2015.8000	2803.3000	2039.7000	2256.2000	2132.2000	1834.1000	1588.0000	1749.4000	1687.200
	2007.7000	2208.3000	1656.7000	1604.4000	1431.8000	1610.5000	1819.4000	2079.7000	2371.6000	2759.900
Fw :	191.5000	516.0000	729.0000	560.4000	519.9000	628.5000	537.1000	561.2000	617.2000	626.700
	737.2000	760.5000	581.4000	662.3000	583.8000	635.2000	723.8000	864.1000	1193.5000	1188.900
Fs :	1362.4000	1807.1000	2676.3000	1801.9000	1957.3000	2202.8999	2380.5000	2168.6000	1985.1000	1813.900
	1850.2000	2067.7000	1796.7000	1625.8000	1667.0000	1667.4000	2269.5000	2159.3999	2031.3000	2115.500
Cg :	97.6000	104.4000	118.0000	156.2000	172.6000	186.6000	220.9000	287.8000	319.9000	321.300
	319.6000	346.0000	456.4000	543.4000	618.3000	647.4000	671.3000	726.1000	800.3000	888.900
Cw :	1.8000	.8000	7.4000	18.1000	23.5000	26.5000	36.2000	60.8000	84.4000	91.200
	92.4000	86.0000	111.1000	130.6000	141.8000	136.7000	129.7000	145.5000	174.8000	213.500
Cs :	53.6000	58.5000	118.1000	260.2000	312.7000	254.2000	261.4000	298.7000	301.8000	279.100
	213.0000	132.6000	264.0000	306.9000	351.1000	357.8000	342.1000	444.2000	623.6000	669.700

Gamma OLS fuer :

$$I_g = -9.9561 \text{ Kon} \quad .0266 \text{ Fg} \quad .1517 \text{ Cg} \quad (X.25)$$

(28.9256) (.0144) (.0237)

st+2 : 660.829 () Fr.-grade : 20 !! () "Sigma"-Werte : Werte in Klammern

$$I_w = -.5094 \text{ Kon} \quad .0529 \text{ Fw} \quad .0924 \text{ Cw} \quad (X.26)$$

(7.3897) (.0145) (.0517)

st+2 : 88.662 () Fr.-grade : 20 !! () "Sigma"-Werte : Werte in Klammern

$$I_s = -4.8159 \text{ Kon} \quad .0174 \text{ Fs} \quad .0390 \text{ Cs} \quad (X.27)$$

(13.6337) (.0068) (.0131)

st+2 : 79.110 () Fr.-grade : 20 !! () "Sigma"-Werte : Werte in Klammern

R = Korrelationsmatrix der Residuen :

1.0000	.7290	.4270
.7290	1.0000	.6098
.4270	.6098	1.0000

Inverse von R :

2.1362	-1.5936	.0597
-1.5936	2.7808	-1.0153
.0597	-1.0153	1.5936

Gamma des SUR-Schaetzers nach 0 Iterationen fuer:

$$I_g = -24.1042 \text{ Kon } .0391 \text{ Fg } .1254 \text{ Cg} \quad (X.28)$$

(26.0352) (.0127) (.0223)

s**2 : 717.220 () Fr.-grade : 20 !! () "Sigma"-Werte : Werte in Klammern

$$I_w = .7001 \text{ Kon } .0586 \text{ Fw } .0337 \text{ Cw} \quad (X.29)$$

(6.4845) (.0119) (.0434)

s**2 : 96.355 () Fr.-grade : 20 !! () "Sigma"-Werte : Werte in Klammern

$$I_s = 1.7922 \text{ Kon } .0149 \text{ Fs } .0337 \text{ Cs} \quad (X.30)$$

(11.8499) (.0059) (.0121)

s**2 : 80.497 () Fr.-grade : 20 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

Ig :	-.8793	-22.8284	-23.1081	-30.7071	-37.7363	-8.3616	37.5488	17.6088	-23.3518	-25.599
Iw :	-1.1036	54.0248	48.9218	39.0719	-11.6580	-27.1073	3.4128	8.4207	9.8430	-6.420
Is :	-.9504	-5.0562	-8.6070	-11.2505	-13.1299	-9.8432	15.1243	7.7131	-2.6832	-2.679
	-7.7329	5.3079	17.0540	5.6572	-7.6415	-10.2808	6.9047	15.5528	13.5676	-8.948
	-2.8745	5.1644	1.4204	-11.1297	-18.4008	-6.9600	1.2764	.4494	-5.3266	-9.355
	-10.6432	5.0209	4.6125	13.1340	2.0909	3.2290	11.4482	15.6597	11.0747	-9.902

R der SURE-Residuen :

1.0000	.7864	.5451
.7864	1.0000	.6985
.5451	.6985	1.0000

Maximale rel. Abweichung : .341566E+00 nach 0 Iterationen

Gamma des SUR-Schaetzers nach 11 Iterationen fuer:

$$I_g = -24.3869 \text{ Kon } .0418 \text{ Fg } .1137 \text{ Cg} \quad (X.31)$$

(26.0120) (.0125) (.0227)

s**2 : 771.560 () Fr.-grade : 20 !! () "Sigma"-Werte : Werte in Klammern

$$I_w = 1.4289 \text{ Kon } .0601 \text{ Fw } .0136 \text{ Cw} \quad (X.32)$$

(6.3843) (.0115) (.0424)

s**2 : 102.939 () Fr.-grade : 20 !! () "Sigma"-Werte : Werte in Klammern

$$I_s = 2.0412 \text{ Kon } .0153 \text{ Fs } .0300 \text{ Cs} \quad (X.33)$$

(10.9706) (.0054) (.0116)

s**2 : 81.621 () Fr.-grade : 20 !! () "Sigma"-Werte : Werte in Klammern

Residuen des SUR-Schaetzers fuer:

Ig :	-2.5886	-26.7886	-29.0709	-34.0758	-41.4948	-11.5998	35.5694	17.1585	-23.8405	-25.898
Iw :	-2.3036	52.5998	50.4219	41.8208	-7.4735	-23.0550	7.1914	12.1760	13.7238	-2.477
Is :	-.0267	-6.5358	-10.2702	-12.4475	-14.1379	-10.9721	14.3268	7.3756	-2.6203	-2.501
	-7.6951	5.1820	17.7002	6.5762	-6.3801	-9.1986	7.7141	16.4721	14.5978	-7.141
	-3.5035	4.3346	.4705	-11.1854	-18.3292	-7.2079	.9796	.3798	-5.3269	-9.347
	-10.6502	4.3827	4.5759	13.3249	2.4270	3.5897	11.4871	16.1397	12.2595	-8.593

R der SURE-Residuen :

1.0000	.8103	.5830
.8103	1.0000	.7199
.5830	.7199	1.0000

Maximale rel. Abweichung : .835270E-05 nach 11 Iterationen

Abb. (X.K)

SURE-Residuen aus Gleichung (X.31)

Regressand war : Ig

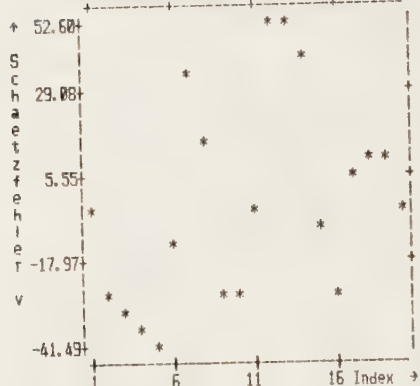
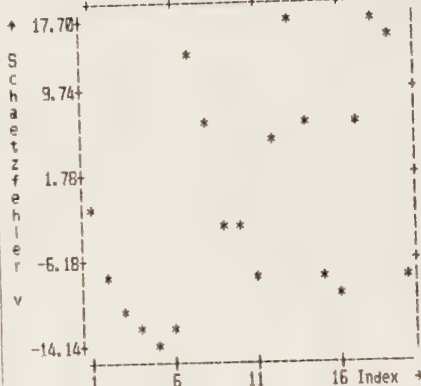


Abb. (X.L)

SURE-Residuen aus Gleichung (X.32)

Regressand war : Iw



Fuer Iw sind Instrumentalvariable: Is Kon Fs Cs

$$I_g = 2.2389 I_w - 37.1660 Kon .0644 F_g -.1965 F_w -.0045 C_g .6065 C_w \quad (X.34)$$

$$(2.9182) (-1.2500) (1.8263) (-1.6059) (-.0559) (1.5319)$$

s**2 : 353.2868 () B(adj.) : .973585 () Fr.-grade : 14 () "t"-Werte : Werte in Klammern

Fuer Iw sind Instrumentalvariable: Is Kon Cs

$$I_g = 2.2259 I_w - 36.9072 Kon .0641 F_g -.1948 F_w -.0041 C_g .6037 C_w \quad (X.35)$$

$$(2.8634) (-1.2327) (1.8010) (-1.5769) (-.0504) (1.5157)$$

s**2 : 356.5837 () B(adj.) : .973338 () Fr.-grade : 14 () "t"-Werte : Werte in Klammern

Fuer Ig sind Instrumentalvariable: Is Kon

$$I_g = 2.6888 I_w - 46.0754 Kon .0774 F_g -.2526 F_w -.0194 C_g .7012 C_w \quad (X.36)$$

$$(4.4851) (-1.8960) (2.6935) (-2.5735) (-.2871) (2.1424)$$

s**2 : 246.7701 () B(adj.) : .981549 () Fr.-grade : 14 () "t"-Werte : Werte in Klammern

Fuer Ig sind Instrumentalvariable: Is Kon Fs Cs

$$I_w = .4289 I_g 16.7311 Kon -.0288 F_g .0892 F_w .0033 C_g -.2685 C_w \quad (X.37)$$

$$(10.1323) (4.8580) (-7.8111) (8.2837) (.3029) (-5.4958)$$

s**2 : 6.3825 () B(adj.) : .997227 () Fr.-grade : 14 () "t"-Werte : Werte in Klammern

Fuer Ig sind Instrumentalvariable: Is Kon Cs

$$I_w = .4463 I_g 16.6024 Kon -.0288 F_g .0878 F_w .0020 C_g -.2708 C_w \quad (X.38)$$

$$(7.5702) (3.7498) (-6.0791) (8.2781) (.1470) (-4.3058)$$

s**2 : 10.5324 () B(adj.) : .995424 () Fr.-grade : 14 () "t"-Werte : Werte in Klammern

Fuer Ig sind Instrumentalvariable: Is Kon

$$I_w = .3722 I_g 17.1491 Kon -.0288 F_g .0940 F_w .0072 C_g -.2610 C_w \quad (X.39)$$

$$(9.8471) (4.7294) (-7.4157) (8.4089) (.6411) (-5.0771)$$

s**2 : 7.0918 () B(adj.) : .996919 () Fr.-grade : 14 () "t"-Werte : Werte in Klammern

Von OLS/EIV verwendete Daten fuer : (t = 1 - 10)

Ig :	33.1000	45.0000	77.2000	44.6000	48.1000	74.4000	113.0000	91.9000	61.3000	56.800
Iw :	12.9300	25.9000	35.8500	22.8900	18.8400	28.5700	48.5100	43.3400	37.0200	37.610
Is :	20.9900	35.5300	46.9900	26.2300	23.0400	36.1600	47.2800	44.5600	36.1600	28.820
Kon :	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.000
Fg :	1170.6000	2015.6000	2803.3000	2039.7000	2256.2000	2132.2000	1834.1000	1588.0000	1749.4000	1687.200
Fw :	191.5000	516.0000	729.0000	560.4000	519.9000	620.5000	537.1000	561.2000	617.2000	626.700
Fs :	1362.4000	1807.1000	2676.3000	1801.9000	1957.3000	2202.8999	2380.5000	2168.6001	1985.1000	1813.900
Cg :	97.8000	104.4000	118.2000	156.2000	172.6000	186.6000	220.9000	287.8000	319.9000	321.300
Cw :	1.8000	.8000	7.4000	18.1000	23.5000	26.5000	36.2000	60.8000	84.4000	91.200
Cs :	53.6000	50.5000	118.1000	260.2000	312.7000	254.2000	261.4000	296.7000	301.6000	279.100

Fuer Iw sind Instrumentalvariable: Is Kon Fs Cs

Ig = 2.1141 Iw -63.9458 Kon .0237 Fg -.0803 Fw .5637 Cg -1.4469 Cw (X.40)
(3.8170) (-1.9010) (.8969) (-.8831) (2.1037) (-2.1726)

s**2 : 87.9707 () B(adj.) : .983196 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern

Fuer Iw sind Instrumentalvariable: Is Kon

Ig = 2.2452 Iw -63.1846 Kon .0261 Fg -.0914 Fw .5318 Cg -1.3639 Cw (X.41)
(4.6634) (-2.2055) (1.1569) (-1.1730) (2.3182) (-2.3911)

s**2 : 63.7871 () B(adj.) : .987816 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern

Fuer Ig sind Instrumentalvariable: Is Kon Fs Cs

Iw = .4383 Ig 27.5979 Kon -.0117 Fg .0414 Fw -.2292 Cg .5877 Cw (X.42)
(8.4674) (3.3762) (-2.2269) (2.3719) (-3.0710) (3.1215)

s**2 : 4.0217 () B(adj.) : .996646 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern

Fuer Ig sind Instrumentalvariable: Is Kon

Iw = .4454 Ig 28.1426 Kon -.0116 Fg .0407 Fw -.2369 Cg .6075 Cw (X.43)
(10.1386) (4.1798) (-2.7016) (2.8486) (-3.8038) (3.8629)

s**2 : 2.6772 () B(adj.) : .997768 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern

Von OLS/EIV verwendete Daten fuer : (t = 11 - 20)

Ig :	93.6000	159.9000	147.2000	146.3000	98.3000	93.5000	135.2000	157.3000	179.5000	189.6000
Iw :	39.2700	53.4600	55.5600	49.5600	32.0400	32.2400	54.3000	71.7000	90.0000	68.6000
Is :	25.8700	42.0300	42.0500	49.4500	40.5100	41.8500	58.8200	64.5500	64.1000	45.9300

Kon :	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Fg :	2007.7000	2208.3000	1656.7000	1504.4000	1431.8000	1610.5000	1819.4000	2079.7000	2371.6000	2759.9000
Fw :	737.2000	760.5000	581.4000	662.3000	583.9000	635.2000	723.8000	864.1000	1193.5000	1188.9000
Fs :	1850.2000	2057.7000	1796.7000	1625.0000	1667.0000	1667.4000	2289.5000	2159.3999	2031.3000	2115.5000
Cg :	319.6000	346.0000	456.4000	543.4000	618.3000	647.4000	671.3000	726.1000	800.3000	888.9000
Cw :	92.4000	86.0000	111.1000	130.6000	141.8000	136.7000	129.7000	145.5000	174.8000	213.5000
Cs :	213.8000	132.6000	264.6000	306.9000	351.1000	357.8000	342.1000	444.2000	623.6000	669.7000

Fuer Iw sind Instrumentalvariable: Is Kon Fs Cs

Ig = 2.3046 Iw -25.9170 Kon .0791 Fg -.2290 Fw -.1001 Cg .9332 Cw (X.44)
(3.4492) (-.7208) (2.1002) (-2.0927) (-1.0078) (1.6768)

s**2 : 211.8711 () B(adj.) : .990788 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern

Fuer Iw sind Instrumentalvariable: Is Kon

Ig = 2.7779 Iw -34.5356 Kon .0857 Fg -.2812 Fw -.1334 Cg 1.1634 Cw (X.45)
(11.8474) (-2.6196) (6.1858) (-7.1243) (-3.6884) (5.7633)

s**2 : 28.7930 () B(adj.) : .999748 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern

Fuer Ig sind Instrumentalvariable: Is Kon Fs Cs

Iw = .4700 Ig 10.6677 Kon -.0360 Fg .0985 Fw .0412 Cg -.3982 Cw (X.46)
(6.7704) (1.6403) (-4.7090) (6.3866) (2.4071) (-4.6006)

s**2 : 7.5977 () B(adj.) : .997924 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern

Fuer Ig sind Instrumentalvariable: Is Kon

Iw = .3600 Ig 12.4325 Kon -.0308 Fg .1012 Fw .0400 Cg -.4188 Cw (X.47)
(3.8435) (.8930) (-1.9793) (3.0521) (1.3235) (-2.2555)

s**2 : 35.4492 () B(adj.) : .990314 () Fr.-grade : 4 () "t"-Werte : Werte in Klammern

Von TSL5 verwendete Daten fuer :

y1 :	2.0000	-.5000	.5000	-.5000	-2.0000	.5000
y2 :	2.0000	-.5000	-.5000	-.5000	1.0000	-1.5000
y3 :	1.0000	-1.5000	-1.5000	-1.5000	0.0000	3.5000
x1 :	-2.0000	-1.0000	0.0000	0.0000	1.0000	2.0000
x2 :	0.0000	-1.0000	1.0000	1.0000	0.0000	-1.0000
x3 :	0.0000	-1.0000	2.0000	0.0000	1.0000	-2.0000

y1 = -1.0000 y2 -1.0000 x1 (-.6245) (-.9874) (XI.1)

s**2 : 2.500 Durbin-Watson :2.300 "Fr.-grade" : 4 "t"-Werte : Werte in Klammern

y2 = -1.0000 y3 -1.0000 x2 (-.8257) (-.5222) (XI.2)

s**2 : 5.500 Durbin-Watson :2.000 "Fr.-grade" : 4 "t"-Werte : Werte in Klammern

y3 = -1.0000 y1 -1.0000 x3 (-.6296) (-1.1495) (XI.3)

s**2 : 7.000 Durbin-Watson :2.536 "Fr.-grade" : 4 "t"-Werte : Werte in Klammern

Pi(i) fuer :

y1 :	-.50000	.50000	-.50000
y2 :	-.50000	-.50000	.50000
y3 :	.50000	-.50000	-.50000

(XI.4)
(XI.5)
(XI.6)

"Prognosen" aus der Strukturform fuer :

y1 :	-.0000	1.5000	.5000	.5000	-2.0000	-.5000
y2 :	-1.0000	2.5000	.5000	.5000	0.0000	-2.5000
y3 :	-2.0000	1.5000	-2.5000	.5000	1.0000	1.5000

"Prognosefehler" aus Strukturform fuer :

y1 :	2.0000	-2.0000	-.0000	-1.0000	.0000	1.0000
y2 :	3.0000	-3.0000	-1.0000	-1.0000	1.0000	1.0000
y3 :	3.0000	-3.0000	1.0000	-2.0000	-1.0000	2.0000

Prognose aus reduzierter Form fuer :

y1 :	1.0000	.5000	-.5000	.5000	-1.0000	-.5000
y2 :	1.0000	.5000	.5000	-.5000	-.0000	-1.5000
y3 :	-1.0000	.5000	-1.5000	-.5000	.0000	2.5000

Prognosefehler aus red. Form fuer :

y1 :	1.0000	-1.0000	1.0000	-1.0000	-1.0000	1.0000
mit Streuung :		2.00000				
y2 :	1.0000	-1.0000	-1.0000	-.0000	1.0000	0.0000
mit Streuung :		1.33333				
y3 :	2.0000	-2.0000	0.0000	-1.0000	-.0000	1.0000
mit Streuung :		3.33333				

$$y1 = .3750 \ y2 \quad (XI.7)$$

(.5590)

s**2 : 1.800 Durbin-Watson :2.279 "Fr.-grade" : 5 "t"-Werte : Werte in Klammern

$$y2 = -.5000 \ y3 \quad (XI.8)$$

(-1.0205)

s**2 : 2.400 Durbin-Watson :1.641 "Fr.-grade" : 5 "t"-Werte : Werte in Klammern

$$y3 = -.5000 \ y1 \quad (XI.9)$$

(-.3764)

s**2 : 5.350 Durbin-Watson :1.409 "Fr.-grade" : 5 "t"-Werte : Werte in Klammern

Pi(i) fuer :

y1 :	0.0000	0.0000	0.0000
y2 :	0.0000	0.0000	0.0000
y3 :	0.0000	0.0000	0.0000

"Prognosen" aus der Strukturform fuer :

y1 :	.7500	-.1875	-.1875	-.1875	.3750	-.5625
y2 :	-.5000	.7500	.7500	.7500	0.0000	-1.7500
y3 :	-1.0000	.7500	-.2500	.2500	1.0000	-.2500

"Prognosefehler" aus Strukturform fuer :

y1 :	1.2500	-.3125	.6875	-.3125	-2.3750	1.0625
y2 :	2.5000	-1.2500	-1.2500	-1.2500	1.0000	.2500
y3 :	2.0000	-1.7500	-1.2500	-1.7500	-1.0000	3.7500

Von TSL5 verwendete Daten fuer :

y1 :	2.0000	-.5000	.5000	-.5000	-2.0000	.5000
y2 :	2.0000	-.5000	-.5000	-.5000	1.0000	-1.5000
y3 :	1.0000	-1.5000	-1.5000	-1.5000	0.0000	3.5000
x2 :	0.0000	-1.0000	1.0000	1.0000	0.0000	-1.0000
x3 :	0.0000	-1.0000	2.0000	0.0000	1.0000	-2.0000

y1 = -.5625 y2
(-.4376)

(XI.10)

s**2 : 2.644 Durbin-Watson :1.429 "Fr.-grade" : 5 "t"-Werte : Werte in Klammern

y2 = -1.0000 y3 -1.0000 x2
(-.4554) (-.3594)

(XI.11)

s**2 : 5.500 Durbin-Watson :2.000 "Fr.-grade" : 4 "t"-Werte : Werte in Klammern

y3 = -1.0000 y1 -1.0000 x3
(-.2315) (-.9449)

(XI.12)

s**2 : 7.000 Durbin-Watson :2.536 "Fr.-grade" : 4 "t"-Werte : Werte in Klammern

Pi(i) fuer :

y1 :	.36000	-.36000
y2 :	-.64000	.64000
y3 :	-.36000	-.64000

(XI.13)

(XI.14)

(XI.15)

y1 = .5000 x2 -.4000 x3
(.4226) (-.5345)

(XI.16)

s**2 : 2.100 Durbin-Watson :1.671 "Fr.-grade" : 4 "t"-Werte : Werte in Klammern

y2 = -.5000 x2 .6000 x3
(-.4841) (.9186)

(XI.17)

s**2 : 1.600 Durbin-Watson :1.475 "Fr.-grade" : 4 "t"-Werte : Werte in Klammern

y3 = -.5000 x2 -.6000 x3
(-.3478) (-.6599)

(XI.18)

s**2 : 3.100 Durbin-Watson :2.116 "Fr.-grade" : 4 "t"-Werte : Werte in Klammern

Pi(i) fuer :

y1 :	.50000	-.40000
y2 :	-.50000	.60000
y3 :	-.50000	-.60000

(XI.19)

(XI.20)

(XI.21)

Von TSLS verwendete Daten fuer :

```

y1 : 2.0000 -1.5000 .5000 -1.5000 -2.0000 .5000
y2 : 2.0000 -1.5000 -1.5000 -1.5000 1.0000 -1.5000
y3 : 1.0000 -1.5000 -1.5000 -1.5000 0.0000 3.5000

x1 : -2.0000 -1.0000 0.0000 0.0000 1.0000 2.0000
x2 : 0.0000 -1.0000 2.0000 0.0000 1.0000 -2.0000

```

y1 = -1.0000 y2 -1.0000 x1 (XI.22)
 (-.4899) (-.8137)

s**2 : 2.500 Durbin-Watson :2.300 "Fr.-grade" : 4 "t"-Werte : Werte in Klammern

y2 = -.5584 y3 (XI.23)
 (-1.0695)

s**2 : 2.624 Durbin-Watson :1.596 "Fr.-grade" : 5 "t"-Werte : Werte in Klammern

y3 = -1.0000 y1 -1.0000 x3 (XI.24)
 (-.5855) (-1.1429)

s**2 : 7.000 Durbin-Watson :2.536 "Fr.-grade" : 4 "t"-Werte : Werte in Klammern

Pi(i) fuer :

y1 : -.64167 -.35833 (XI.25)
 y2 : -.35833 .35833 (XI.26)
 y3 : .64167 -.64167 (XI.27)

y1 = -.5000 x1 -.2500 x3 (XI.28)
 (-1.2271) (-.6136)

s**2 : 1.594 Durbin-Watson :2.824 "Fr.-grade" : 4 "t"-Werte : Werte in Klammern

y2 = -.5000 x1 .2500 x3 (XI.29)
 (-1.4813) (.7407)

s**2 : 1.034 Durbin-Watson :1.829 "Fr.-grade" : 4 "t"-Werte : Werte in Klammern

y3 = .5000 x1 -.7500 x3 (XI.30)
 (.9619) (-1.4429)

s**2 : 2.594 Durbin-Watson :2.217 "Fr.-grade" : 4 "t"-Werte : Werte in Klammern

Pi(i) fuer :

y1 : -.50000 -.25000 (XI.31)
 y2 : -.50000 .25000 (XI.32)
 y3 : .50000 -.75000 (XI.33)

Von OLS* verwendete Daten fuer :

```

y1 : 9.0000 5.0000 -1.0000 -3.0000 -5.0000 -5.0000 4.0000 -4.0000
y2 : 4.0000 7.0000 -3.0000 -5.0000 -3.0000 1.0000 7.0000 -8.0000

x1 : -2.0000 -1.0000 0.0000 0.0000 1.0000 2.0000 0.0000 0.0000
x2 : 0.0000 -1.0000 1.0000 1.0000 0.0000 -1.0000 -2.0000 2.0000

```

y1 = .5000 y2 -3.0000 x1 (XII.1)
(8.0911) (-10.3035)

s**2 : .750 () d : 2.778 () B(adj.) : .9735 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 1, 5 Freiheitsgraden : F* : .00

y2 = .5000 y1 -3.0000 x2 (XII.2)
(7.4087) (-10.9434)

s**2 : .750 () d : 2.278 () B(adj.) : .9764 () Fr.-grade : 6 () "t"-Werte : Werte in Klammern
Joint-Test auf Unabhaengigkeit mit 1, 5 Freiheitsgraden : F* : .00

Beta-Invers fuer :

```

y1 : 1.3333 .6667
y2 : .6667 1.3333

```

Pi(i) fuer :

```

y1 : -4.0000 -2.0000 (XII.3)
y2 : -2.0000 -4.0000 (XII.4)

```

Korrelationsmatrix der U :

```

1.0000 -.5000
-.5000 1.0000 !!

```

Korrelationsmatrix der V :

```

1.0000 .5000
.5000 1.0000

```

Korrelationsmatrix der V und U :
→ Zeilen = V → Spalten = U :

```

.8660 -.0000 !!
-.0000 .8660

```

Prognosen aus Strukturform fuer :

```

y1 : 8.0000 6.5000 -2.5000 -4.5000 -5.5000 3.5000 -4.0000
y2 : 4.5000 5.5000 -3.5000 -4.5000 -2.5000 .5000 8.0000 -8.0000

```

Prognosefehler aus Strukturform (=U) fuer :

```

y1 : 1.0000 -1.5000 .5000 -.5000 -.5000 .5000 .5000 .0000
y2 : -.5000 1.5000 .5000 -.5000 -.5000 .5000 -1.0000 0.0000

```

Prognose aus reduzierter Form fuer :

```

y1 : 8.0000 6.0000 -2.0000 -2.0000 -4.0000 -6.0000 4.0000 -4.0000
y2 : 4.0000 6.0000 -4.0000 -4.0000 -2.0000 -.0000 8.0000 -8.0000

```

Prognosefehler aus red. Form (=V) fuer :

```

y1 : 1.0000 -1.0000 1.0000 -1.0000 -1.0000 1.0000 -.0000 .0000
mit Streuung : 1.0000

y2 : -.0000 1.0000 1.0000 -1.0000 -1.0000 1.0000 -1.0000 .0000
mit Streuung : 1.0000

```

Von OLS* verwendete Daten fuer :

```

y1 : 5.0000 3.0000 -1.0000 -3.0000 -3.0000 -1.0000 4.0000 -4.0000
y2 : 4.0000 5.0000 -1.0000 -3.0000 -3.0000 -1.0000 3.0000 -4.0000

x1 : -2.0000 -1.0000 0.0000 0.0000 1.0000 2.0000 0.0000 0.0000
x2 : 0.0000 -1.0000 1.0000 1.0000 0.0000 -1.0000 -2.0000 2.0000

```

y1 = .5000 y2 -1.0000 x1 -1.0000 x2 (XII.5)
(1.2910) (-1.2032) (-1.2166)

s**2 : .900 () d : 2.778 () B(adi.) : .9267 () Fr.-grade : 5 () "t"-Werte : Werte in Klammern

y2 = .5000 y1 -1.0000 x1 -1.0000 x2 (XII.6)
(1.2910) (-1.2032) (-1.2166)

s**2 : .900 () d : 2.278 () B(adi.) : .9267 () Fr.-grade : 5 () "t"-Werte : Werte in Klammern

Beta-Invers fuer :

```

y1 : 1.3333 .6667
y2 : .6667 1.3333

```

Pi(i) fuer :

```

y1 : -2.0000 -2.0000
y2 : -2.0000 -2.0000

```

(XII.7)
(XII.8)

Korrelationsmatrix der U :

```

1.0000 -.5000
-.5000 1.0000      II

```

Korrelationsmatrix der V :

```

1.0000 .5000
.5000 1.0000

```

Korrelationsmatrix der V und U

→ Zeilen = V + Spalten = U :

```

.8660 .0000
.0000 .8660      II

```

Prognosen aus Strukturform fuer :

```

y1 : 4.0000 4.5000 -1.5000 -2.5000 -2.5000 -1.5000 3.5000 -4.0000
y2 : 4.5000 3.5000 -1.5000 -2.5000 -2.5000 -1.5000 4.0000 -4.0000

```

Prognosefehler aus Strukturform (=U) fuer :

```

y1 : 1.0000 -1.5000 .5000 -.5000 -.5000 .5000 -.5000 -.0000
y2 : -.5000 1.5000 .5000 -.5000 -.5000 .5000 -1.0000 -.0000

```

Prognose aus reduzierter Form fuer :

```

y1 : 4.0000 4.0000 -2.0000 -2.0000 -2.0000 -2.0000 4.0000 -4.0000
y2 : 4.0000 4.0000 -2.0000 -2.0000 -2.0000 -2.0000 4.0000 -4.0000

```

Prognosefehler aus red. Form (=V) fuer :

```

y1 : 1.0000 -1.0000 1.0000 -1.0000 -1.0000 1.0000 .0000 -.0000
mit Streuung : 1.0000

```

```

y2 : .0000 1.0000 1.0000 -1.0000 -1.0000 1.0000 -1.0000 -.0000
mit Streuung : 1.0000

```

L I T E R A T U R V E R Z E I C H N I S

=====

- (1) Aigner D. : " MSE dominance of least squares with errors-of-observation " ; Journal of Econometrics 2 (1974) S. 365 - 372
- (2) Bartlett M.S. : " Fitting a straight line when both variables are subject to error " ; Biometrics 5 (1949) S. 207 - 212
- (3) Basman R. : " On finite sample distributions of generalized classical linear identifiability test statistics " ; Journal of the American Statistical Association 55 (1960) S. 650 - 659
- (4) Boot J., de Witt G. : " Investment demand: An empirical contribution to the aggregation problem " ; International Economic Review , 1 (1960) S. 3 - 30
- (5) Bronstein I., Semendjajew K. : " Taschenbuch der Mathematik " ; Frankfurt - Zuerich 1967
- (6) Christ C. : " Econometric models and methods " ; New York 1966
- (7) Dhrymes P.J. : " Econometrics " ; New York 1970
- (8) Dhrymes P.J. : " Distributed laast: Problems of estimation and formulation " ; San Francisco 1971
- (9) Doran H.E., Griffiths W.E. : " Inconsistency of the OLS estimator of the partial adjustment - adaptive expectations modell " ; Journal of Econometrics 7 (1978) S. 133 - 146
- (10) Durbin J. : " Errors in variables " ; Revue de l'Institut International de Statistique , 22 (1954) S. 23 - 32
- (11) Durbin J. : " Testing for serial correlation in least squares regression when some of the regressors are lagged dependent variables " ; Econometrica 38 , 1970 S. 410 - 421
- (12) Esenwein-Rothe I. : " Theorie der Wirtschaftsstatistik " ; Kompendium der Volkswirtschaftslehre, Band 2; Goettingen 1972 S. 413 - 512
- (13) Fisher F.M. : " On the cost of approximate specification in simultaneous equation estimation " ; Econometrica 29 (1961) S. 139 - 170
- (14) Fisher F.M. : " The identification problem in econometrics " ; New York 1966
- (15) Friedman M. : " A theory of the consumption function " ; Princeton 1957
- (16) Gabrielsen A. : " Consistency and Identifiability " ; Journal of Econometrics 8 , 1978 S. 261 - 263
- (17) Griliches Zwi : " Errors in variables and other unobservables " ; Econometrica 42 (1974) S. 971 - 998
- (18) Gruber G. : " Zur Interdependenz oekonometrischer Gleichungssysteme " ; Unveroeffentlichter Vortrag, gehalten auf der Tagung des "Ausschusses fuer Empirische Wirtschaftsforschung und Angewandte Oekonometrie" der Deutschen Statistischen Gesellschaft, Augsburg 31.5 + 1.6. 1977
- (19) Grunfeld Y. : " The determinants of corporate investment " ; Unveroeffentlichte Ph. D. Thesis , Chicago (1958) , zitiert nach Boot & de Witt (1960)

- (20) Guttman L. : " Multiple rectilinear prediction and the resolution into components " ; Psychometrika 5 (1940) S. 75 - 99
- (21) Guttman L. : " Simple proofs of relations between the communality problem and multiple correlation " ; Psychometrika 22 (1957) S. 147 - 157
- (22) Intriligator M.D. : " Econometric Models, Techniques and Applications " ; Amsterdam - Oxford 1978
- (23) Johnston J. : " Econometric Methods " ; New York 1963
- (24) Johnston J. : " Econometric Methods " ; New York 1972
- (25) Kakwani C.N. : " The unbiasedness of Zellner's seemingly unrelated regression equations estimators " ; Journal of the American Statistical Association, 62 (1967) S. 141 - 142
- (26) Kmenta J., Gilbert R. : " Small sample properties of alternative estimators of seemingly unrelated regression equations estimators " ; Journal of the American Statistical Association, 63 (1968) S. 1180 - 1200
- (27) Koutsoyiannis A. : " Theorie of Econometrics " ; London 1973
- (28) Leamer E.E. : " Specification searches " ; New York 1978
- (29) Leonhard W. : " Einfuehrung in die Regelungstechnik " ; Braunschweig 1972
- (30) Levi M.D. : " Errors in the variables bias in the presence of correctly measured variables " ; Econometrica 41 (1973) S. 985 - 986
- (31) Levi M.D. : " Measurement errors and bounded OLS estimates " ; Journal of Econometrics 6, (1977) S. 165 - 171
- (32) Liu Ta-Chung : " Underidentification, structural estimation, and forecasting " ; Econometrica 28 (1960) S. 855 - 865
- (33) Lyttkens E. : " The iterative instrumental variable method and the full information maximum likelihood method for estimating interdependent systems " ; Journal of Multivariate Analysis (1974) s. 283 - 307 ; zitiert nach Maddala (1977) S. 490
- (34) Madansky A. : " The fitting of straight lines when both variables are subject to errors " ; American Statistical Association Journal 54 (1959) S. 173 - 205
- (35) Maddala G.S. : " Econometrics " ; New York 1977
- (36) Malinvaud E. : " Statistical Methods of Econometrics " ; Amsterdam - London 1970
- (37) McCullum B.T. : " Relative asymptotic bias from errors of omission and measurement " ; Econometrica 40 (1972) S. 757 - 758
- (38) Neumann M. : " Theoretische Volkswirtschaftslehre " ; Heidelberg 1973
- (39) Neyman J., Scott E.L. : " On certain methods of estimating the linear structural relation " ; Annals of Mathematical Statistics, 22 (1951) S. 352 - 361 und Corrections ; Annals of Mathematical Statistics (1952) S. 115

- (40) Novak E. : " Konsistente Schaetzung von Parametern in autoregressiven Systemen mit Fehlern in den Variablen, bei gegebener Kovarianzmatrix der Fehler" ; Forschungsberichte aus dem Institut fuer Statistik und Wissenschaftstheorie der Universitaet Muenchen , Serie De. Nr. 1 , 1975
- (41) Novak E. : " Konsistente Schaetzung saentlicher Parameter einer Zeitreihenregression bei Fehlern in den Variablen, wenn neben den Daten der Zeitreihe keine zusaetzlichen Kenntnisse verfuesbar sind" ; Forschungsberichte aus dem Institut fuer Statistik und Wissenschaftstheorie der Universitaet Muenchen , Serie De Nr. 3 , 1975
- (42) Novak E. : " Identifikation oekonometrischer Zeitreihen-Regressionen bei Fehlern in den Variablen, ohne Benutzung zusaetzlicher Kenntnisse" ; Forschungsberichte aus dem Institut fuer Statistik und Wissenschaftstheorie der Universitaet Muenchen, Serie De. Nr. 6 , April 1976
- (43) Novak E. : " Identifikation von stochastischen Modellen der Zeitreihenanalyse bei Fehlern in den Variablen" ; Forschungsberichte aus dem Institut fuer Statistik und Wissenschaftstheorie der Universitaet Muenchen, Serie De. Nr. 8, 1977
- (44) Ramsey B. : " Tests for specification errors in classical linear least-squares regression analysis" ; Journal of the Royal Statistical Society, Ser. B, 31 (1969) S. 351 - 371
- (45) Rao P. : " Some notes on misspecification in multiple regressions" ; The American Statistician 25 (1971) S. 37 - 39
- (46) Rao P. : " Specification bias in seemingly unrelated regressions" ; in: "Econometrics and Economic Theorie" ; Essays in Honour of J. Tinbergen (Ed.: Sellekaerts) London (1974) S. 101 - 113
- (47) Reiersol D. : " Confluence analysis by means of instrumental sets of variables" ; Arkiv for Matematik, Astronomi Och Fysik 32 (1945) S. 1 - 119 ; Zitiert nach Madansky (1959) S. 205
- (48) Schmidt P. : " A note on the estimation of seemingly unrelated regression systems" ; Journal of Econometrics 7 (1978) S. 259 - 261
- (49) Schneeweiss H. : " Dekonometrie" ; Wuerzburg - Wien 1974
- (50) Schoenfeld P. : " Methoden der Dekonometrie", Band I ; Muenchen 1969
- (51) Schoenfeld P. : " Methoden der Dekonometrie", Band II ; Muenchen 1971
- (52) Seibt H. : " Schaetzung und Prognose bei simultanen Gleichungsmodellen" ; Goettingen 1968
- (53) Stobbe A. : " Wirtschaftskreislauf und Sozialprodukt" ; Kompendium der Volkswirtschaftslehre Band I, Goettingen 1972 S. 16 - 56
- (54) Strotz H. : " Interdependence as a specification error" ; Econometrica 28 (1960) S. 428 - 442
- (55) Strotz H., Wold H.O.A. : " Recursive versus nonrecursive systems: An attempt at synthesis" ; Econometrica 28 (1960) S. 417 - 427
- (56) Telser L. : " Iterative estimation of a set of linear regression equations" ; Journal of the American Statistical Association 59 (1964) S. 842 - 862
- (57) Theil H. : " Economic forecasts and policy" ; Amsterdam 1961
- (58) Theil H. : " Principles of Econometrics" ; Amsterdam - London 1971

- (59) Wald A. : " Fitting of straight lines if both variables are subject to errors " ; *Annals of Mathematical Statistics* 11; (1940) S. 284 - 300
- (60) Weede E., Jagodzinsky W. : " Einfuehrung in die konfirmatorische Faktorenanalyse " ; *Zeitschrift fuer Soziologie* Jg. 6 Heft 3 (1977) S. 315 - 333
- (61) Wickens M.R. : " A note on the use of proxy variables " ; *Econometrica* 40 (1972) S. 759 - 761
- (62) Wold H., Faxer R. : " On the specification error in regression analysis " ; *Annals of Mathematical Statistics*, March 1957
- (63) Wold H.O.A. : " A generalization of causal chain models " ; *Econometrica* 28 (1960) S. 443 - 463
- (64) Wonnacott R.J., Wonnacott Th.H. : " *Econometrics* " ; New York 1970
- (65) Wu De-Min : " Alternative tests of independence between stochastic regressors and disturbances " ; *Econometrica* 41 No. 4, 1973 S. 733 - 750
- (66) Zellner A. : " An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias " ; *Journal of the American Statistical Association*, 57 (1962) S. 348 - 368
- (67) Zellner A., Huang D. : " Further properties of efficient estimators for seemingly unrelated regression equations " ; *International Economic Review* 3 (1962) S. 300 - 313
- (68) Zellner A. : " Estimators for seemingly unrelated regression equations: some exact finite results " ; *Journal of the American Statistical Association*, 58 (1963) S. 977 - 992

Tabellarischer Lebenslauf

Nachname : D r e f a h l

Vorname : D i e t e r

Geburtsdatum : 14. November 1946

Geburtsort : Badenweiler/Krs. Müllheim (Baden-Württemberg)

Name und Beruf : Dr. jur. Werner Drefahl ; Senatspräsident a. D.
des Vaters

Name der Mutter: Lieselotte, geborene Schmitt

Ausbildungsgang: Volksschule 1953 - 1957 in Lörrach und Stuttgart
Gymnasium 1957 - 1966 in Stuttgart und
Nürtingen; Abitur am 9. März 1966
Technisches Industriepraktikum bei IBM Sindelfingen April bis September 1966
Physikstudium an der Universität Stuttgart von
WS 1966/67 bis WS 1968/69
Werkstudent mit Programmierlehrgang bei Siemens
AG Stuttgart April bis August 1969
Programmierlehrgang beim DGB Stuttgart Mai bis
Oktober 1969
Banklehre bei der Deutschen Bank AG Stuttgart
von September 1969 bis Juni 1971; Kaufmanns-
gehilfenprüfung am 28.6.1971
Studium der Betriebswirtschaft an der Universität
Erlangen-Nürnberg von WS 1971/72 bis SS 1975
Studentische Hilfskraft und Tutor am Lehrstuhl
für Statistik II (Prof. Dr. Schneeberger) von
WS 1972/73 bis SS 1975
Diplomprüfung zum Diplomkaufmann am 18.11.1975
Seit 20.11.1975 Verwalter der Dienstgeschäfte
eines wissenschaftlichen Assistenten am Lehr-
stuhl für Statistik II (Prof. Dr. Schneeberger)
am Fachbereich Wirtschafts- und Sozialwissen-
schaften der Friedrich-Alexander-Universität
Erlangen-Nürnberg in Nürnberg
Promotion zum Dr. rer. pol. am 15.7.1980

Druck: Höfer & Limmert, Erlangen